

MESG
MESTRADO EM ENGENHARIA
DE SERVIÇOS E GESTÃO

LAURIUS

**Desenvolvimento de uma estrutura de dados sobre o
comportamento macro dos clientes portadores de um programa de
fidelização**

Laura Isabel Serrão Ferreira

Dissertação de Mestrado

Orientador na FEUP: Prof.^a Eng.^a Teresa Galvão

Orientador na empresa: Dr.^a Liliana Bernardino



2017-06-30

Aos meus pais e irmão

Resumo

Ao longo dos anos, influenciado pela concorrência sentida, o sector do retalho tem crescido não só ao nível de parque de lojas, mas também no sentido do aumento da “exigência” do consumidor. O apoio das novas tecnologias e metodologias permitem que haja uma melhor gestão de informação possibilitando assim o crescimento e sucesso. Segundo Einstein, o tempo é relativo, no entanto é também crucial para que a produtividade das empresas seja notória. Numa empresa que aposta na evolução permanente, na satisfação do cliente e no contacto com este, poderá através de uma estrutura de dados desenvolvida, analisar dados acionáveis do utente do programa de fidelização face ao seu comportamento registado.

A finalidade deste projeto passa por dar apoio à empresa de retalho na eficácia da utilização dos dados recolhidos do cliente portador do programa de fidelização realizando futuramente análises dos indicadores calculados desde segmentações até valores em vendas, viabilizando assim o apoio à tomada de decisão sustentada em momentos decisivos de satisfação deste.

A metodologia utilizada no desenvolvimento da estrutura de dados consiste em analisar os processos anteriormente desenvolvidos e desta forma responder à identificação de necessidades ou requisitos das várias equipas envolvidas na gestão do programa de fidelização. Necessidades essas, algumas anteriormente calculadas para um período definido, outras semelhantes mas com um período diferente de cálculo e outras em forma de sugestão para novas ações promocionais e análises de dados. Em todo o percurso foram aprofundados conhecimentos, em particular a nível de sistemas de informação e as suas relações, foi definida a estratégia de intervenção e levantamento de necessidades, selecionados os indicadores relevantes para as várias áreas de decisão da empresa como o departamento de marketing, programa de fidelização, equipas digitais, bem como áreas de análise comercial e operacional, posteriormente foi planeado o desenvolvimento da estrutura e iniciou-se a elaboração do protótipo da ferramenta cuja necessidade identificada.

Com a criação desta ferramenta este estudo demonstra que o tempo despendido para as ações promocionais e para futuras análises diminuiu, tanto no cálculo das métricas como no tratamento dos dados para futuras ações, análises e decisões promocionais integrando diversas necessidades que são regularmente solicitadas possibilitando assim análises mais detalhadas e objetivas. No entanto apenas foram desenvolvidos alguns dos indicadores tendo em conta dois fatores inicialmente definidos: o facto do período determinado para a realização da dissertação ser curto para a finalização do projeto e o facto de alguns desses indicadores se apresentarem como novos indicadores sugeridos, o que necessitam de ser discutidos a nível de termos e critérios internos da empresa. Todavia, por se mostrar um projeto ambicioso e de grande utilidade para a empresa, futuramente os indicadores não produzidos serão calculados dando assim continuidade e culminação do projeto.

Development of a data structure on the macro behavior loyalty program customers

Abstract

Over the years, influenced by perceived competition, the retail sector has grown not only at the level of retail stores, but also in the direction of increasing consumer "demand". The support of new technologies and methodologies allows for better information management, thus enabling growth and success. According to Einstein, time is relative, yet it is also crucial for the productivity of companies to be noticeable. In a company that bets on the permanent evolution, in the customer satisfaction and contact with the customer, it will be able through a developed data structure, to analyze actionable data of the user of the loyalty program in face of its registered behavior.

The purpose of this project is to support the retail company in the effectiveness of the use of the data collected from the customer holding the loyalty program, conducting in the future analyzes of the indicators calculated from segmentations to values in sales, therefore enabling support for sustained decision-making in crucial moments of customer satisfaction.

The methodology used in the development of the data structure consists of analyzing the previously developed processes and in this way respond to the identification of needs or requirements of the several teams involved in the management of the loyalty program. Needs that were previously calculated for a defined period, others similar but with a different period of calculation and others in a form of a suggestion for new promotional actions and data analysis. Throughout the course were deepened knowledge, in particular at the level of information systems and their relationships, was defined the intervention strategy and needs assessment, selected the relevant indicators for the various areas of decision of the company as the marketing department, loyalty program, digital teams, as well as areas of commercial and operational analysis, later the development of the structure was planned and the prototype of the tool whose need was identified was started.

With the creation of this tool, this study demonstrates that the time spent with the analyzes has decreased, both in the calculation of metrics and in the treatment of the data for future actions, analyzes and promotional decisions, integrating several needs that are regularly requested, thus enabling more detailed and objective analyzes. However, only some of the indicators were developed, taking into account two factors initially defined: the fact that the period determined for the dissertation was short for project conclusion, and the fact that some of these indicators presented themselves as new suggested indicators which need to be discussed at internal company terms and criteria level. Nevertheless, because it proves to be an ambitious project and of great utility for the company, in the future the indicators not produced will be calculated, giving continuity and culmination to the project.

Agradecimentos

Apresento aqui os meus agradecimentos a todas as pessoas que de uma forma ou de outra me ajudaram ao longo deste projeto de dissertação, contribuindo para o sucesso do mesmo.

Agradeço à empresa pelas condições e oportunidade de estágio oferecidas para a realização deste projeto.

A todas as pessoas da empresa com quem tive a oportunidade de trabalhar, em particular à minha orientadora, Dr.^a Liliana Bernardino, que conduziu a minha intervenção no sentido de uma participação pró-ativa, mas que, com disponibilidade e oportunidade, sempre apoiou em alturas chave.

A todos os colaboradores das equipas da direção do programa de fidelização que disponibilizaram a sua atenção e fundamento, em especial à equipa Customer Intelligence and Analytics, Ana Freitas, Patrícia Castro, Ana Januário, Lisandra Rocha, Paulo Castro, Leandro Pinto, Sílvia Cunha, José Ferreira, José Rodrigues, Diana Matos, Ana Ferreira, Liliana Martins, Paulo Granja, Ana Pinto e Rui Carneiro e ainda à Maria Céu Silva pelo apoio, conselhos e forma como ajudaram na integração da equipa, a disponibilidade e colaboração sempre que precisei, e, sem esquecer, a permanente boa energia e boa disposição.

À Faculdade de Engenharia da Universidade do Porto a possibilidade de realizar este projeto para que houvesse contacto com o ambiente empresarial e, assim, facilitar a integração no mundo de trabalho.

Agradeço à diretora de curso Lia Patrício pela sua dedicação e empenho em tornar o Mestrado em Engenharia de Serviços e Gestão num excelente e desafiante mestrado com muito boas aprendizagens e referências.

À Prof.^a. Teresa Galvão, pelo apoio e incentivo oferecidos, ajuda e atenção na elaboração da dissertação nesta etapa final do mestrado.

Aos meus Pais e ao meu Irmão por todo o amor incondicional, pelo apoio, incentivo, preocupação, pelo sacrifício de estarem longe para que eu pudesse chegar até aqui e tornar-me na pessoa que sou.

À minha família e amigos por todo o apoio e ajuda incondicional, em especial à família Martins.

Peculiarmente, agradecer à Patrícia Silva por todo o ensinamento, apoio e camaradagem ao longo desta jornada.

Por último mas não com menos importância, um obrigado especial ao Fábio Martins pelo apoio constante, pelo companheirismo, pelas palavras de ânimo, segurança e força transmitida em todas as horas e por não me ter deixado desistir.

Índice de Conteúdos

1	Introdução.....	1
1.1	Contexto do projeto.....	1
1.2	Descrição do problema.....	1
1.3	Objetivos de investigação.....	2
1.4	Organização da Dissertação.....	2
2	Estado da arte.....	4
2.1	Marketing no sector do Retalho Alimentar.....	4
2.2	Big Data e Customer Behavior.....	5
2.3	Business Intelligence.....	6
2.4	Sistemas de Informação.....	7
2.5	Bases de dados relacionais.....	7
2.6	SAS - Statistical Analysis System Enterprise Guide.....	8
2.7	Particionamento.....	8
3	Desenvolvimento do projeto.....	9
3.1	Apresentação da Empresa.....	9
3.2	Oportunidade levantada pela empresa.....	9
3.3	Objetivos para a empresa.....	11
3.4	Objetivos de investigação.....	11
3.5	Parceiros do programa de fidelização.....	12
3.6	Indicadores da estrutura de dados.....	12
4	Metodologia.....	14
4.1	Análise comparativa de abordagens existentes e fundamentação da escolha da abordagem adotada.....	14
4.2	Método seguido no projeto.....	17
4.2.1	Definição do problema.....	17
4.2.2	Levantamento de Requisitos.....	17
4.2.3	Exploração de ferramentas, abordagens e metodologias.....	21
4.2.4	Implementação.....	22
5	Resultados.....	27
6	Conclusões e perspectivas de trabalho futuro.....	33
	Referências.....	35
	Bibilografia.....	37
ANEXO A:	Planeamento do Projeto LAURIUS.....	38
ANEXO B:	198 Indicadores recolhidos na fase de Levantamento de requisitos.....	39
ANEXO C:	Documento de Levantamento de requisitos.....	59
ANEXO D:	Indicadores da estrutura de dados LAURIUS.....	60
ANEXO E:	Modelo de dados geral do projeto LAURIUS.....	74
ANEXO F:	Modelo de dados detalhado do projeto LAURIUS.....	75

ANEXO G:	Tabelas utilizadas dos Schemas lab e mkt e da View E, que pertencem à base de dados analítica.	76
ANEXO H:	Tabelas temporárias desenvolvidas para o cálculo dos indicadores.	78
ANEXO I:	Tabelas das partições implementadas.	79
ANEXO J:	Amostra tabela LAURIUS	81

Lista de Figuras

Figura 1 - Definição da oportunidade levantada.....	10
Figura 2 – Grupo de indicadores de segmentação.....	19
Figura 3 - Grupo de indicadores de contacto com o cliente	19
Figura 4 – Grupo de Indicadores de informação transacional	20
Figura 5 - Grupo de indicadores de informação do cliente	20
Figura 6 – Grupo de indicadores com informação promocional	21
Figura 7 – Grupo de indicadores com informação adicional	21
Figura 8 – Esquematização da metodologia de Agregação/Particionamento Vertical	23
Figura 9 – Exemplo da metodologia aplicada na tabela LAURIUS	24
Figura 10 – Ligação ternária entre o Cliente, Parceiro e Tempo	24
Figura 11 – Exemplo do <i>Schema</i> lab do Anexo G.....	25
Figura 12 – Exemplo da <i>View</i> E do Anexo G	25
Figura 13 – Exemplo do Anexo H com algumas tabelas temporárias criadas	26
Figura 14 – Exemplo do Anexo I com algumas tabelas do particionamento para posterior agregação	26
Figura 15 – Coluna com a data de atualização da tabela LAURIUS	26
Figura 16 – Tabela criada com a base de clientes filtrados em SAS	27
Figura 17 – Indicadores a roxo em fase de avaliação de critérios internos e a rosa indicadores ‘fora da caixa’	27
Figura 18 – Indicadores implementados a azul	28
Figura 19 - Canais de contacto direto ao cliente	29
Figura 20 – Hora preferencial de compra no último mês nas lojas do parceiro Q	30
Figura 21 – Segmentação 4 dos clientes com loja preferencial 10 do parceiro A	31
Figura 22 – Percentagem de vendas brutas na segmentação 4 dos clientes do parceiro A	32
Figura 23 – Percentagem de vendas líquidas na DC Casa do parceiro A nos últimos 12 meses	32

Lista de siglas e abreviaturas

1FN - Primeira Forma Normal

2FN - Segunda Forma Normal

3FN - Terceira Forma Normal

Ad hoc – solução designada para uma tarefa ou problema específico

BI – Business Intelligence

CCR - Customer Centric Retail

CTT - Correios de Portugal (Correios, Telégrafos e Telefones)

DC – Direção Comercial

EDW - Enterprise Data Warehouse

IT - Information Technology

IVA - Imposto sobre o Valor Acrescentado

LAURIUS –Loyalty program, Actionable, Updated, Requirements, Insight, User view, Storage

LEDB - Loyalty Engine Database

MDM – Master Data Management

SAS - Statistical Analysis System

SQL - Structured Query Language

U12M – Últimos 12 meses

U24M - Últimos 24 meses

Umes – Último mês

UN - Unidade

VB – Vendas Brutas

VL – Vendas Líquidas

1 Introdução

No âmbito da Dissertação realizada no Mestrado em Engenharia de Serviços e Gestão, o presente documento resulta de uma proposta lançada pela empresa portuguesa SONAE à Faculdade de Engenharia da Universidade do Porto. Esta proposta insere-se num Programa de Estágios para Novos Desafios, criado pela empresa cujo objetivo consiste em aliar o conhecimento académico ao conhecimento da mesma. Esta empresa tem vindo assim a dominar o mercado por apresentar conceitos e procedimentos especializados sobre os recursos da empresa com vista à resolução de desafios reais por si enfrentados, promovendo-se, assim, a difusão de ideias e do conhecimento em ambos os sentidos.

Por motivos de confidencialidade e de restrição de acesso, determinadas informações não serão mencionadas neste documento.

1.1 Contexto do projeto

Nos dias que correm existe, por parte das empresas, uma necessidade crescente em lidar com muita informação pelo que é crucial que existam novas técnicas e metodologias que facilitem a gestão e a sincronização dos dados, bem como a resolução de incoerências ou inconsistências. Numa empresa de retalho, o contacto com o cliente é cada vez mais importante e considerado um dos grandes objetivos a cumprir. A análise de dados referentes ao cliente exige que haja uma maior eficácia na resposta à ativação de comunicação em determinado momento para o programa de fidelização em questão.

A finalidade do projeto apresentado vai de encontro a estas necessidades que, até à data, existia apenas de uma forma isolada e individual, sendo efetuadas análises específicas consoante o objetivo pretendido. Através do desenvolvimento de uma estrutura de dados de alimentação automática e que detenha informação relevante sobre o comportamento (macro) do cliente detentor do programa de fidelização, será possível tomar decisões mais objetivas apoiadas no foco do cliente.

1.2 Descrição do problema

O período de ação de contacto com o cliente pode parecer relativo. No entanto, é essencial para que as decisões tomadas tenham sucesso dando assim resposta à eficiência e ao tempo exigidos nas decisões tomadas. As equipas responsáveis pela gestão do programa de fidelização precisam de uma estrutura de dados que dê apoio nesse sentido, pois a informação e os indicadores existentes encontram-se calculados de forma isolada, o que impossibilita um acesso rápido e eficaz nas análises e ações futuras.

A proposta de projeto visa então criar uma estrutura de dados que guarde informação relevante sobre o comportamento do cliente detentor do programa de fidelização criado pela empresa de retalho. Esta estrutura de dados é de alimentação automática e capaz de responder às necessidades base de uma futura decisão mais objetiva, sendo um grande apoio em decisões com o foco no cliente.

Os três grandes objetivos deste projeto, integrado no negócio de retalho alimentar, são os seguintes: a nível de desenvolvimento, pretende-se produzir um repositório de informação atualizada; a nível de utilidade, pretende-se garantir que todos os indicadores definidos estejam em conformidade com as necessidades das equipas envolvidas e, por fim, a nível de contacto, pretende-se com este projeto melhorar a comunicação com o cliente.

A estrutura de dados criada possibilita uma automatização e organização de informação, crucial para fomentar o interesse do cliente no seu programa de fidelização. O seu desenvolvimento cruza a análise e a gestão de bases de dados, sendo necessário cruzar a base de dados operacional Loyalty Engine Database (LEDB) com a base de dados analítica Enterprise Data Warehouse (EDW) já existentes na empresa.

1.3 Objetivos de investigação

O negócio de retalho está em crescente competitividade no mercado, exigindo que sejam estabelecidas estratégias de vantagem competitiva para otimizar os processos indispensáveis para aumentar o lucro da empresa e o interesse dos clientes.

O principal objetivo passa por construir, utilizando a metodologia apropriada, uma estrutura de dados suportada por uma arquitetura Master Data Management (MDM), com informação organizada e fiável acerca do comportamento do cliente (Customer Behavior) face ao programa de fidelização. As diversas variáveis calculadas permitem que dados das bases de dados operacionais e analíticas sejam cruzados para que posteriormente se possam realizar análises desses mesmos dados, explorando assim a área de Business Intelligence.

Outro objetivo de investigação passa por reunir todos os indicadores identificados como necessários pelas várias equipas envolvidas na gestão do programa de fidelização. Posteriormente, na fase de seleção dos indicadores finais, pretende-se incluir na estrutura de dados apenas aqueles que têm como base os objetivos comuns das diferentes equipas, de acordo com a sua prioridade e periodicidade (Diária/Semanal). Desta forma, cumpre-se o objetivo de reduzir o número total de indicadores essenciais, criando assim uma estrutura de dados com carácter de grande utilidade para a empresa.

Desenvolver um modelo de dados relacional coerente com as ligações existentes entre os dados e tabelas é ainda outro objetivo de investigação, permitindo às equipas visualizar e compreender melhor o conceito de sistema de informação implementado.

Através da estrutura desenvolvida com todos os dados necessários, será possível tomar decisões mais sustentadas no que respeita ao foco no cliente e nas suas opções de escolha disponibilizadas pelo programa de fidelização. Este acesso rápido aos dados essenciais permitirá reduzir o tempo necessário para desenvolver documentos como análises e relatórios, influenciando assim a eficácia de qualquer processo a este nível.

A garantia da qualidade e coerência dos dados explorados e calculados é também outro objetivo de investigação. A validação dos dados utilizados e calculados devem ter em conta conceitos como a integridade, concisão e consistência a fim de minimizar a margem de erro em conclusões de posteriores análises.

Para além da qualidade dos dados, outro objetivo é assegurar a segurança dos mesmos. Os dados utilizados devem estar protegidos tendo em conta a sua confidencialidade.

1.4 Organização da Dissertação

Este documento encontra-se dividido em seis capítulos: Introdução; Estado da arte; Desenvolvimento do projeto; Metodologia; Resultados; Conclusões e perspectivas de trabalho futuro.

No primeiro capítulo é feita uma introdução, na qual é abordado o âmbito do projeto desenvolvido, bem como o contexto do mesmo. Em seguida é referida a descrição do

problema e os objetivos de investigação delineados para o projeto e, finalmente, a descrição de como se encontra organizado o presente documento.

No segundo capítulo é apresentada a revisão de literatura referente às áreas de Marketing em relação ao sector de Retalho Alimentar referindo as áreas de Big Data, Customer Behavior e Business Intelligence. Apresentam-se também conceitos de Sistemas de Informação, Bases de dados relacionais, a ferramenta SAS e a metodologia de particionamento de tabelas em bases de dados.

No terceiro capítulo é abordado o desenvolvimento do projeto, iniciando pela apresentação da empresa em que o projeto foi desenvolvido. Em seguida, apresenta-se o desafio lançado, quais os seus objetivos apresentando os parceiros do programa de fidelização e, por fim, analisa-se a importância e os critérios de seleção dos indicadores da estrutura de dados desenvolvida.

O quarto capítulo é dedicado à metodologia utilizada nas várias fases do desenvolvimento do projeto, apresentando também as metodologias semelhantes encontradas.

O quinto capítulo é dedicado aos resultados obtidos com a nova estrutura de dados e à análise da utilidade dos indicadores calculados com a apresentação de alguns gráficos.

No sexto capítulo são apresentadas as conclusões retiradas do desenvolvimento do projeto e dos resultados obtidos, bem como propostas para trabalhos futuros e melhorias do mesmo.

2 Estado da arte

Neste capítulo é apresentada a literatura mais relevante para o projeto nas áreas de Marketing em relação ao sector Retalho Alimentar, referindo as áreas de Big Data, Customer Behavior e Business Intelligence. Definem-se ainda alguns conceitos de Sistemas de Informação, Bases de dados relacionais, apresenta-se a ferramenta SAS e a metodologia de particionamento de tabelas em bases de dados.

Inicialmente, de modo a compreender a área em que o projeto foi desenvolvido, são apresentados alguns conceitos inerentes ao Marketing no Retalho Alimentar e depois, mais direcionado ao projeto, as atividades que na literatura estão definidas como pertencentes ao Customer Behavior, Business Intelligence e Sistemas de Informação.

Posteriormente, é apresentada uma metodologia encontrada a propósito do desenvolvimento do projeto referindo as suas vantagens e inconvenientes.

2.1 Marketing no sector do Retalho Alimentar

Estudar e analisar o ambiente que envolve o Marketing permite às empresas aproveitar novas oportunidades de vantagem competitiva para combater possíveis ameaças. Além disso, a gestão do Marketing pode também ser afetada por atores ou outras forças externas ao Marketing, o que pode prejudicar relações de sucesso com clientes futuros.

O estudo do Marketing pode resumir-se à informação utilizada para identificar e definir oportunidades e problemas, gerindo, melhorando e avaliando as ações tomadas, controlando a performance do marketing e progredindo assim na sua compreensão (Stoicescu, 2016).

Por conseguinte, reunir informação está a tornar-se progressivamente importante no Marketing, uma vez que se verifica um aumento da competição, não só a nível de vendas, mas também a nível de informação. O Marketing está a tornar-se uma batalha baseada mais em informação do que no poder de vendas (Kotler, Armstrong, Saunders, & Wong, 1999).

Marketing é muitas vezes associado ao termo vender. Segundo a definição no dicionário, vender significa existir uma transação entre duas partes onde o comprador recebe bens (tangíveis ou intangíveis) ou serviços em troca de dinheiro.

No entanto, cada vez mais o conceito de Marketing não significa apenas vender, mas também satisfazer as necessidades do cliente. Mais do que qualquer outra função de negócios, Marketing lida com os clientes. A criação de valor e satisfação do cliente está no centro da prática e do pensamento do marketing moderno (Kotler, Wong, Saunders, & Armstrong, 2005). Toda a gestão de Marketing passa por analisar, planear, implementar e controlar todas as operações com esse fim, aumentando a escolha disponível para o cliente em quantidade e qualidade.

A criação de valor do cliente define-se como a satisfação por existir o produto ou serviço que pretende. Quando um cliente precisa de satisfazer uma necessidade, procura um produto ou um serviço com as características que lhe são mais satisfatórias, segundo as características que esses produtos ou serviços lhe oferecem. A escolha final é parcialmente ligada à qualidade pois esta tem um impacto direto no desempenho do produto (Kotler et al., 2005). Se o cliente considerar que o produto ou serviço que comprou é de qualidade significa que o desempenho desse produto ou serviço estava de acordo com as expectativas de satisfação que o cliente trazia ou ainda que foram superadas.

Kotler (Kotler et al., 2005) refere ainda que o retalho é definido por todas as atividades de venda de bens ou serviços diretamente aos consumidores finais. Assim, no sector do retalho alimentar existe uma variedade de produtos expostos em que o cliente pode escolher por forma a satisfazer as suas necessidades. Necessidades essas que foram pensadas e estrategicamente planeadas segundo a informação recolhida, para que seja captada a informação do cliente final, fazendo-o comprar o produto na loja. O principal requisito nestes sectores é extrair informação útil dos dados recolhidos (Kocakoç & Erdem, 2010).

2.2 Big Data e Customer Behavior

Embora pareça que o mercado esteja numa fase de amadurecimento e saturação, há sempre uma forma de inovar com novos conceitos de negócio, com o objetivo de atingir bons resultados locais e globais. Existe uma necessidade de antever as tendências e expectativas dos clientes desenvolvendo e implementando soluções que os satisfaçam. Desta forma, a empresa pode ganhar vantagem sobre os seus concorrentes, assim como a atenção e fidelização por parte dos clientes.

O cliente evoluiu nos conceitos de poupança de tempo e dinheiro no sentido em que a tecnologia permitiu que este tivesse melhores apoios e ferramentas para estar informado sobre todos os produtos e serviços à sua disposição. Segundo Stoicescu (Stoicescu, 2016), o processo de compra de um determinado produto simplificou-se em termos de tempo e dinheiro gasto, no entanto tornou-se mais difícil e complexo o processo de decisão. Isto deve-se ao facto das pessoas terem várias opções de escolha, em termos não só de produto ou serviço, mas também de preço, qualidade e tempo.

Numa empresa de retalho, todo o processo de compra do cliente é considerado. Através do programa de fidelização associado às decisões de negócio tomadas, o comportamento do cliente fica registado. Esta informação registada em grande volume (Big Data), é seguidamente gerida para que as empresas consigam compreender e ponderar as próximas ações.

As ferramentas de análise de dados são perfeitas para atingir resultados exatos e aumentar o lucro (Stoicescu, 2016). Na realidade, a análise de Big Data possibilita a extração de conhecimento para encontrar a definição de novos conceitos, padrões camuflados ou até relações inesperadas que podem fazer a diferença no mundo do Marketing. Para explorar e compreender o comportamento do cliente, o conhecimento extraído ajuda na identificação das suas necessidades. As empresas perceberam que a análise do comportamento do cliente é um fenómeno crescente que evoluiu com o desenvolvimento humano, tornando-se interessante estudar o comportamento diário do cliente (Stoicescu, 2016). Como resultado, as decisões tomadas pelas empresas são influenciadas pelo valor instantâneo que o cliente “exige”.

Stoicescu (Stoicescu, 2016) também concluiu que são vários os principais fatores que formam o comportamento do cliente. Entre estes fatores contam-se as alterações demográficas no crescimento da população e das expectativas de vida que resultam na satisfação de mais necessidades, a evolução da tecnologia que permite que as pessoas estejam informadas sobre as novas culturas e estilos de vida que adotam ao viajar, originando assim novas e mais diversificadas necessidades, a evolução dos meios de pagamento, a eficiência na rapidez e no custo da satisfação das necessidades e, por fim, o risco e o stress por haver muitas opções de escolha para satisfazer as próprias necessidades.

O conhecimento sobre o cliente é uma necessidade importante e diferenciadora, pois torna possível prever o comportamento deste em decisões futuras.

2.3 Business Intelligence

Business Intelligence (BI) está integrado na área de Marketing que gradualmente vai desenvolvendo a sua importância nas áreas de estudo dos seus praticantes e investigadores, refletindo a magnitude e o impacto dos problemas relacionados com dados nas organizações empresariais contemporâneas (Chen, Chiang, & Storey, 2012). As empresas têm estimulado uma melhoria nos recursos que têm disponíveis para conseguir efetivamente investir na tecnologia da informação.

Segundo Stackowiak (Stackowiak, Rayman, & Greenwald, 2007), esta área é o processo de utilizar grandes quantidades de dados, analisá-los e apresentar um conjunto de relatórios de alto nível que resumem a essência daqueles dados em ações de negócio base, permitindo que a equipa de gestão tome diariamente decisões fundamentais de negócio.

BI, para Golfarelli, Rizzi e Cella (Golfarelli, Rizzi, & Cella, 2004), pode ser definido como o processo de transformar dados em informação e depois em conhecimento que é tipicamente obtido sobre as necessidades do cliente, os processos de influência na decisão do cliente, a competição, as condições do negócio e as tendências económicas, tecnológicas e culturais em geral.

Para Ranjan (Ranjan, 2009) BI é composto por alguns elementos como OLAP (On-line analytical processing), que se refere à forma como os utilizadores do negócio podem dividir e usar os dados utilizando ferramentas sofisticadas, que permitem a navegação em dimensões como o tempo e a hierarquia. Posteriormente, este elemento fornece visões multidimensionais e resumidas de dados de negócios usadas para relatórios, análises, criação de modelos e planeamento de otimização do negócio. Advanced Analytics é o segundo componente referido, relacionado com técnicas como data mining e análises de previsão que usam técnicas de análise estatística para prever ou fornecer determinadas medidas ou factos. A Gestão de Desempenho Corporativo, baseada em scorecards e dashboards, é um componente que possibilita agregar vários dados, que depois parecem contar uma história dos factos de forma a retirar conhecimento e evoluir nas métricas calculadas.

Por último, o componente de BI que Ranjan (Ranjan, 2009) considera mais significativo é o Data Warehouse. Data Warehouse significa armazém dos dados, é orientado e integrado no sentido em que tem a capacidade de sustentar a propagação física dos dados, gerindo os numerosos registos da empresa para integração, limpeza, agregação e tarefas de consultas das bases de dados. Este armazém de dados tem como fonte bases de dados que podem ser operacionais ou relacionais. As operacionais possuem histórico de dados, dados externos ou informação sobre o ambiente do armazém de dados existente. As relacionais suportam as aplicações de negócio utilizadas.

É nesta área que as empresas alcançam os objetivos de extrair informação útil através dos dados recolhidos. BI permite a utilização de estratégias e tecnologias de gestão de dados que suportam as análises de dados e a apresentação da informação das empresas sustentando assim uma série de decisões eficientes (Dedić & Stanier, 2016).

A extração de informação gera a grande oportunidade de aumentar o conhecimento sobre as necessidades do cliente. Esta é uma necessidade encontrada pelas empresas que exige investigação, de forma a melhorar as decisões tomadas direcionadas para este. A gestão dos dados recolhidos pode ser feita pela inovação nos sistemas de informação que se reveste de particular importância. A presença e relevância destes é cada vez maior, pois suportam

operações consideravelmente mais complexas, dando a possibilidade de implementar novos conceitos e oferecer novos serviços ao cliente (Kocakoçy, 2010).

2.4 Sistemas de Informação

Segundo Alter (Alter, 2002), um sistema de informação é um sistema de *software* capaz de capturar, transmitir, armazenar, recuperar, manipular e apresentar informação de modo a apoiar pessoas, organizações ou outros sistemas de software.

Um sistema de informação ajuda na transformação de dados em informação, permitindo a gestão destes e é composto por pessoas, *hardware*, *software*, base de dados, programas de aplicação e processos. Completando a definição de um sistema de informação, Coronel e Morris (Coronel & Morris, 2016) afirmam que basicamente uma base de dados é um repositório de factos que é cuidadosamente desenhado e construído, fazendo parte de um todo que constitui o sistema de informação.

Portanto, os principais componentes de um sistema de informação são a informação recolhida, as atividades ou ações realizadas, a tecnologia utilizada, produtos ou serviços usados ou desenvolvidos e as pessoas envolvidas no desenvolvimento incluindo os próprios utilizadores deste.

2.5 Bases de dados relacionais

Os dados recolhidos são então guardados numa base de dados, que se define como sendo uma estrutura partilhada e integrada, que permite armazenar dados. Essa estrutura é composta por classes (tabelas) com um nome atribuído a cada uma, constituídas por registos individuais (linhas) e campos/atributos (colunas) que são comuns a todos os registos individuais armazenados. Cada tabela deve representar uma única questão e várias instâncias da mesma. Os campos ou atributos devem ser simples, com apenas um possível resultado de retorno e não devem ser redundantes, salvo para manter histórico, apresentação ou exatidão da transação.

No desenho do modelo relacional composto por tabelas, atributos e relações da base de dados, alguns dos atributos são designados por chaves. As chaves consistem em um ou mais atributos que determinam outros atributos e podem ser de vários tipos. A chave primária é um conceito importante porque assegura que cada linha da tabela é única. Para que haja integridade nos dados e nas tabelas, as chaves primárias têm de ter valores únicos e não pode ter um valor nulo. O valor nulo (*null*) é a inexistência de um valor de um atributo.

Outro tipo de chave é a chave estrangeira. Este atributo, presente numa dada tabela, tem valores que se referem aos valores da chave primária de uma outra tabela permitindo assim relacionar duas tabelas distintas através de um atributo comum. Através delas, é possível relacionar as tabelas criadas garantindo a integridade dos dados (Date, 1999).

A chave secundária é uma chave que se define como uma chave simplesmente usada para fins de recuperação de dados. Em particular, Coronel e Morris (Coronel & Morris, 2016) mostram um exemplo sobre os dados de um cliente armazenados numa tabela A com a chave primária “Número cliente”. Na maioria dos casos, é difícil um cliente memorizar o seu número de cliente. A recuperação de dados para o cliente torna-se mais fácil quando se usa o seu último nome e o número de telefone. Nesse caso, a chave primária será o número do cliente e a chave secundária é a combinação do último nome do cliente com o número de telefone. No entanto, uma chave secundária não retorna necessariamente um resultado único.

2.6 SAS - Statistical Analysis System Enterprise Guide

SAS é um *software* desenvolvido pelo instituto multinacional Americano SAS (SAS, 2013) caracterizado por desenvolver *software* analítico. Dirigido para as áreas de Advanced Analytics, Business Intelligence, Gestão de Dados e Análise Preditiva, este *software* fornece um ambiente autónomo de análise IT (Information Technology), segurança centralizada baseada em funções e gerida pela IT, permite fácil acesso às bases de dados das empresas para diferentes utilizadores e distribui conteúdo interativo para utilizadores do Microsoft Office e Web.

Como componentes, o SAS tem bibliotecas que são espaços físicos, onde são guardadas as tabelas usadas por este. Permite transformar todas as fontes de informação em bibliotecas, nomeadamente pastas do Windows ou esquemas de Oracle e permite o cruzamento entre todas as bibliotecas dentro de um projeto. Existem dois tipos de bibliotecas: as bibliotecas Work, que correspondem ao espaço temporário guardado apenas durante o intervalo de tempo daquela sessão de SAS e as bibliotecas User correspondentes ao espaço criado por cada utilizador que guarda as tabelas durante o tempo desejado.

Para além da linguagem com os comandos próprios para criar e transformar tabelas (DATA), selecionar amostras (PROC SURVEYSELECT), entre outros, o SAS também possibilita utilizar código em SQL (Structured Query Language), facilitando assim a utilização deste *software* na gestão dos dados.

As macros são funções criadas para chamar um bloco de código com algumas consultas (queries) que podem ser chamadas por um comando fácil reaproveitando assim o código. São importantes para a criação de procedimentos automáticos à semelhança das macros em Excel (Microsoft Office) e para a execução de ciclos à semelhança dos cursores em Oracle.

2.7 Particionamento

Os sistemas de informação nas empresas de retalho gerem dados que por sua vez são utilizados para uma série de operações. O volume de dados por vezes é grande, o que pode dificultar a gestão dos mesmos. O particionamento de um sistema de informação é vantajoso para o desempenho e para a simplificação da sua manutenção. Esta solução permite que uma tabela grande seja dividida em pequenas tabelas individuais melhorando assim o tempo de execução das consultas de dados.

Existem três tipos de particionamento ou fragmentação: Horizontal, Vertical e Mixed (Coronel & Morris, 2016). A fragmentação horizontal possibilita ter uma tabela dividida em subtabelas com linhas únicas mantendo as colunas. A fragmentação vertical permite que numa base de dados distribuída a tabela se divida em pequenas tabelas em que as linhas se mantêm e as colunas são diferentes da tabela original, mas mantendo a chave primária da tabela inicial. A fragmentação mixed resulta na combinação da horizontal e da vertical, onde a tabela é dividida em várias linhas e cada uma tem colunas diferentes. No entanto, a nível de performance e integridade apresentam-se como desvantagens desta metodologia.

3 Desenvolvimento do projeto

Neste capítulo é apresentada a empresa onde o projeto foi desenvolvido, bem como o problema levantado, explicando com mais detalhe os objetivos para a organização e os objetivos de investigação. Na última parte é apresentada uma breve descrição dos parceiros do programa de fidelização e a importância dos indicadores selecionados para a estrutura de dados final.

3.1 Apresentação da Empresa

Este projeto foi desenvolvido ao longo de quatro meses (de Fevereiro a Junho de 2017) numa das maiores empresas nacionais de retalho. A missão desta empresa é criar valor social e económico, procurando levar os benefícios do progresso à comunidade, atuando com base na ética e confiança, colocando as pessoas no centro do sucesso investindo no desenvolvimento das suas capacidades e competências, ter ambição, inovação e responsabilidade social, otimizar a utilização de recursos, maximizando o seu retorno, mantendo um perfil de simplicidade e eficiência, cooperando e permanecendo independente.

Possui parcerias estratégicas nos negócios dos centros comerciais e das telecomunicações, tendo ainda uma área de gestão do património imobiliário de retalho e outra de gestão de investimentos ativos.

Esta empresa integra um programa de fidelização ao nível do retalho alimentar com parcerias internas e externas. O negócio de retalho alimentar e de retalho especializado destacam-se como os dois negócios fundamentais da empresa.

O negócio de retalho alimentar encontra-se organizado em várias marcas com o objetivo de alcançar os consumidores finais através de diferentes superfícies, desde hipermercados até supermercados de conveniência. O negócio de retalho especializado inclui marcas que abrangem áreas da eletrónica, desporto e moda.

3.2 Oportunidade levantada pela empresa

Gerir informação engloba processar, manipular e organizar dados que futuramente geram conhecimento em diferentes perspetivas dependendo do utilizador desta. A concorrência a nível do conhecimento e satisfação do cliente no mundo empresarial do retalho necessita cada vez mais de informação relevante nesse sentido.

Renovar os conceitos de negócio é uma perspetiva que suporta as empresas nas fases sazonais e que parecem não evoluir. O objetivo da empresa de querer antever as tendências e expectativas dos seus clientes pode ser pensado e cumprido através de soluções que mantenham também a fidelização dos mesmos.

Esta investigação contribui assim para avançar na área de gestão de dados dos clientes detentores de um programa de fidelização criado pela empresa. A solução passa por desenvolver uma ferramenta estruturada de dados que permite suportar decisões futuras em termos de ações para satisfazer o cliente final, organizando de forma acionável os seus dados.

De forma a responder automaticamente e com exatidão às exigências do consumidor e cliente do programa de fidelização criado, a empresa enfrenta o desafio de criar e gerir informação a vários níveis. Os dados existentes encontram-se guardados e tratados, mas de forma individual. Neste sentido, o desenvolvimento de uma ferramenta que detenha todos os indicadores necessários para agir em conformidade com o comportamento macro do cliente

participante no cartão do programa de fidelização, será um grande apoio nas decisões associadas a todo o tipo de campanhas promocionais e análises de dados, não só no sentido de conseguir gerir o grande volume de dados de um modo organizado, mas também de utilizar esses dados para gerar e obter informação crucial, mantendo o interesse do cliente e aumentando o crescimento da empresa.

O cruzamento de dados é realizado com o apoio das bases de dados analítica e operacional existentes na empresa, Enterprise Data Warehouse (EDW) e Loyalty Engine Database (LEDB) respetivamente. A Figura 1 ilustra de uma forma geral o conceito do projeto.

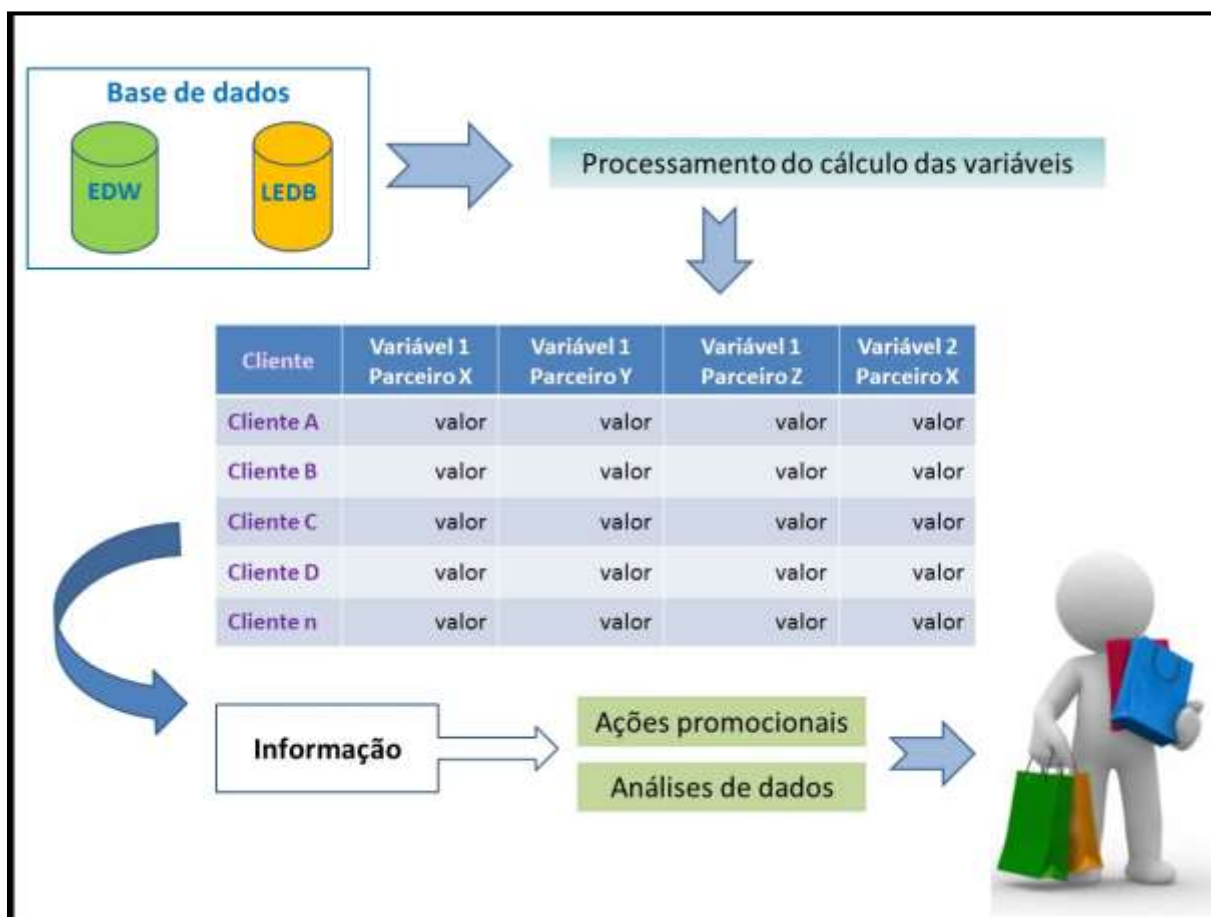


Figura 1 - Definição da oportunidade levantada

Este projeto dividiu-se em várias fases, como a de levantamento de necessidades, análise e extração de dados, aplicação da metodologia de implementação e desenho da ferramenta, apresentadas e explicadas com detalhe no capítulo 4.

Cumprindo o objetivo da oportunidade levantada pela empresa, o conhecimento sobre o cliente é uma necessidade importante e diferenciadora no mercado, que requer uma atenção maior para que haja vantagem competitiva na criação de valor para o cliente consumidor final. Cruzando e utilizando os dados das bases de dados existentes, é possível calcular os indicadores identificados pelas equipas gestoras do programa de fidelização, resultando numa tabela com todos esses indicadores numa perspetiva do cliente. Com base na tabela desenvolvida, é gerada informação decisiva nas próximas ações tomadas para satisfazer e melhorar a comunicação com o cliente.

3.3 Objetivos para a empresa

A empresa definiu como objetivo principal o desenvolvimento de uma estrutura de dados que possibilitasse uma automatização e organização de informação crucial para fomentar o interesse do cliente no seu programa de fidelização, bem como, para melhorar a comunicação com este. Esta estrutura reúne todos os indicadores calculados de modo a ser possível ter uma melhor perspetiva, obtendo assim mais conhecimento sobre o cliente.

Analisando com mais detalhe os objetivos da organização, foi proposto que:

- a estrutura de dados deve conter dados do programa de fidelização (**Loyalty program**),
- os dados devem ser acionáveis para que possam ser utilizados a qualquer momento (**Actionable**),
- os dados devem estar atualizados não existindo histórico (**Updated**),
- os indicadores calculados são baseados nos requisitos e necessidades das equipas envolvidas na gestão do programa de fidelização (**Requirements**),
- futuramente serão realizadas análises de forma a compreender melhor os dados (**Insight**),
- cada linha da tabela representa um cliente detentor do programa de fidelização (**User view**)
- e todos os dados resultantes do cálculo dos indicadores encontram-se armazenados numa única estrutura de dados (**Storage**).

Extraíndo as iniciais de cada objetivo determinado, resulta na palavra LAURIUS sendo este o nome do projeto desenvolvido.

Para além destes, a empresa definiu ainda que os clientes envolvidos são aqueles que se encontram ativos no programa de fidelização, considerando que não devem existir dados idênticos.

Para a empresa, a noção de cliente individual corresponde uma família. Assim procurando compreender e conhecer melhor os seus clientes, e conseguindo assim prever futuras opções destes, a empresa pretende uma estrutura de dados relacionados com o comportamento dos mesmos, ajudando e dando acesso a descontos ou promoções que satisfaçam uma família.

3.4 Objetivos de investigação

O processo metódico de construção do conhecimento em relação à metodologia utilizada na criação da estrutura de dados pretendida, abordou diversas finalidades. A metodologia utilizada é apresentada no próximo capítulo 4, descrevendo com mais detalhe cada uma das fases definidas. No Anexo A é possível verificar o planeamento da investigação e do projeto realizado ao longo dos 4 meses.

Segundo Pressman (Pressman, 2005), o levantamento de necessidades e requisitos aparenta à partida ser um processo simples. No entanto, perceber quais são realmente os requisitos importantes é uma das fases mais difíceis de resolução de um problema. Neste projeto, o período de compreensão de levantamento de necessidades é dos mais importantes, pois é a base da estrutura final. Um dos objetivos de investigação é entender de forma clara e simples, os indicadores mais importantes identificados pelas equipas envolvidas na gestão do programa de fidelização.

Na fase seguinte, que corresponde a selecionar os requisitos que irão fazer parte da ferramenta desenvolvida, os critérios utilizados têm como objetivos: garantir que cada requisito é único, eliminando os repetidos; considerar como necessidade comum de diferentes equipas; definir a prioridade com que precisam do indicador e a usabilidade deste em termos semanais ou mensais. Através destes critérios, o número de requisitos levantados diminuiu para um número significativo.

O desenvolvimento de um modelo de dados ajuda a que os utilizadores da ferramenta implementada consigam entender melhor como os dados se relacionam na base de dados. O modelo desenvolvido é coerente com as ligações e associações existentes, apesar de, por motivos de confidencialidade e também por não ser relevante, o modelo original do sistema de informação atualmente implementado não será aqui apresentado.

Atendendo aos objetivos da ferramenta criada, são utilizados grandes volumes de dados, o que exige uma gestão eficaz e assertiva destes aquando do cálculo dos indicadores finais. Wedel e Kannan (Wedel & Kannan, 2016) afirmam que Big Data é normalmente caracterizado pelos quatro V's: volume, velocidade, variedade e veracidade. Os dois primeiros são importantes de um ponto de vista computacional e os restantes são importantes de um ponto de vista analítico. Outro objetivo sobre os dados é a coerência e segurança dos mesmos. Os indicadores calculados devem ter lógica no sentido em que a sua definição não se altera independentemente do período de cálculo, permitindo assim que as equipas envolvidas detenham todas a mesma definição. Tratando-se de um programa de fidelização que agrega dados dos seus utilizadores, a segurança e confidencialidade dos dados deverá ser necessariamente assegurada na ferramenta final.

3.5 Parceiros do programa de fidelização

O programa de fidelização criado pela empresa reúne não só vantagens para os seus clientes, utentes do cartão, mas também para os parceiros associados. As parcerias aumentam as vantagens para o cliente, que poderá assim, utilizar o cartão do programa de fidelização nas lojas dos parceiros. As regalias para os parceiros são a prestação de um serviço e a gestão de fundos do programa de fidelização, aumentando também o número de clientes em ambos os sentidos. Existem iniciativas que fomentam este esforço crescente do ecossistema gerando assim a frequência entre os diferentes parceiros.

Por motivos de exigências de confidencialidade exigidas pelos parceiros, é apenas possível afirmar que foram considerados na estrutura de dados desenvolvida um total de vinte e um parceiros de diferentes negócios e setores como alimentar, moda, saúde ou energia. Cada um deles foi substituído por uma letra do alfabeto, de A a U, sendo que os indicadores calculados nem sempre consideraram os diferentes parceiros por não ser possível o cálculo ou por não ter sido um dos requisitos levantados pelas equipas.

3.6 Indicadores da estrutura de dados

Atendendo aos objetivos da empresa, a identificação dos indicadores são uma parte importante da realização do projeto pois são a base da estrutura desenvolvida. Tendo em conta a dimensão do documento, todos os indicadores recolhidos encontram-se no Anexo B. Estes indicadores foram retirados do documento de levantamento de requisitos criado, num total de 198.

Dos 198 indicadores recolhidos, são apenas apresentados os oitenta e três indicadores selecionados, assim como uma breve descrição de cada um deles, incluindo a sua definição, tipo de retorno, período de atualização e utilidade. Os oitenta e três indicadores selecionados

são o resultado da fase de levantamento de requisitos e estão divididos em dois grupos: os que já se encontram no sistema de informação atual ou já foram calculados, e os que não existem. Para os indicadores que não existem: alguns podem ser calculados reutilizando código de implementação; outros que finalizaram o processo de avaliação de critérios internos, tendo agora condições para que sejam implementados; outros ainda encontram-se em fase de avaliação de critérios internos e, por fim, existem indicadores nunca antes considerados que pertencem a um conceito de diferenciação designado por 'fora da caixa'.

4 Metodologia

Neste capítulo é apresentada a metodologia aplicada ao longo do projeto desenvolvido. A primeira secção indica a metodologia escolhida entre as várias abordagens encontradas na pesquisa realizada e a segunda secção apresenta o método levado a cabo, bem como os diferentes passos e procedimentos.

4.1 Análise comparativa de abordagens existentes e fundamentação da escolha da abordagem adotada

Durante a pesquisa realizada, foram consideradas e analisadas algumas abordagens para as diferentes fases do desenvolvimento do projeto.

Para os quatro meses estabelecidos de desenvolvimento do projeto, a ferramenta desenvolvida foi planeada e organizada em quatro grandes fases: Definição do problema; Levantamento de requisitos; Exploração e análise das ferramentas, abordagens e metodologias e, por último, Implementação da ferramenta LAURIUS com base na metodologia selecionada.

Fase 1: Definição do Problema

Na fase de Definição do Problema, a metodologia utilizada foi a comunicação e o diálogo em reuniões com os orientadores do projeto de forma a esclarecer os objetivos do mesmo, bem como a indicação da importância da aprendizagem dos modelos existentes de negócio e do modelo de dados do programa de fidelização. Inicialmente houve um período de exploração e compreensão das tarefas e capacidades dos elementos das equipas, sendo estes os utilizadores finais da ferramenta desenvolvida. Segundo Evaluation Research Team (Evaluation Research Team, 2008), observar os utilizadores no seu meio de trabalho ajuda na compreensão de algumas necessidades que por vezes estão subentendidas. Ainda que este método consuma algum tempo, para além de reunir os requisitos das equipas, é necessário compreender de que forma os dados são usados e geridos, sem que elas o digam ou expliquem como o fazem.

Fase 2: Levantamento de requisitos

Outro método interessante e utilizado na fase seguinte de reunir informação e necessidades para o projeto, foram as entrevistas às equipas. Segundo uma pesquisa feita (Evaluation Research Team, 2009), este método pode ser útil para reunir dados qualitativos e quantitativos que sejam válidos e fidedignos. Dos três tipos de entrevistas (Informal de conversação, Semiestruturada e Padrão), a Semiestruturada foi o tipo de entrevista mais adequado. A entrevista abrangeu pontos e perguntas chave, porém não existiu nenhuma ordem padrão das perguntas e sim uma conversa fluente com esclarecimento de dúvidas caso necessário. Este método é muito útil para a compreensão e contexto do assunto em questão e permite ainda ao utilizador da ferramenta referir que elementos são importantes para ele, juntando assim frases ou palavras essenciais para cumprir os objetivos do projeto. Tal como o método da observação, a entrevista consome algum tempo e pode ser interpretada pelos entrevistados como uma técnica intrusiva.

Ainda na fase de levantamento de requisitos, e também pensando na fase de implementação da estrutura de dados, foram exploradas as metodologias avaliadas por List, Bruckner, Machaczek e Schiefer (List, Bruckner, Machaczek, & Schiefer, 2002). Estes avaliaram e resumiram em três grupos metodologias diferentes no desenvolvimento de um armazém de dados (*data warehouse*) sendo eles *data-driven*, *goal-driven* e *user-driven*. A abordagem

data-driven é recomendada para a área de data mining e exploração de dados com foco no cliente e é frequentemente criticada pelo tempo exigido, o que requer um acompanhamento mais firme. A metodologia *goal-driven* refere-se a uma metodologia de suporte a métodos de gestão moderna e é a base das decisões de apoio a toda a organização. O seu tempo de desenvolvimento é extenso e dispendioso e os utilizadores da ferramenta são raramente envolvidos. Em relação à abordagem *user-driven*, caracteriza-se por ser de certa forma arriscada pois reflete o nível organizacional das pessoas envolvidas. Por vezes, pode tornar-se uma visão egocêntrica no sentido em que as pessoas envolvidas não conseguem abstrair e pensar de um modo mais amplo na organização. Esta metodologia deve ser combinada com uma das duas anteriores de forma a aumentar a longevidade do sistema.

Assim, a fase de Levantamento de Requisitos foi dividida em três etapas: a própria fase de levantamento de requisitos em conjunto com os utilizadores da ferramenta, a análise dos resultados recolhidos e a seleção de variáveis. A metodologia utilizada nesta fase é uma combinação dos três grupos mencionados por List, Bruckner, Machaczek, e Schiefer (List et al., 2002) pois envolve as áreas de data mining, foco no cliente, gestão de objetivos no suporte de decisões e entrevistas conduzidas para reunir requisitos dos utilizadores da ferramenta. Esta fase é muito importante para reunir as necessidades das equipas integradas na gestão do programa de fidelização porque os requisitos recolhidos são a base da tabela final construída. Por conseguinte, os elementos destas equipas são os utilizadores finais da estrutura de dados LAURIUS, que utilizaram os dados desta tabela para análises de data mining e outras ações.

A informação recolhida ao longo da observação e das entrevistas foi registada num documento de requisitos. Para Smith (Smith, 2011), um documento de requisitos é imprescindível porque explica o contexto da necessidade de um produto, contendo a lista das necessidades dos seus utilizadores. Este documento facilita o processo de seleção final dos requisitos, visto que se encontram classificados conforme as especificações definidas. Alterando um pouco a estrutura definida por Smith (Smith, 2011) mas mantendo o intuito da metodologia, o documento de requisitos teve como base uma tabela em que cada linha é um requisito que respeita as várias especificações, tais como:

- Número do requisito,
- Nome da equipa,
- Nome do indicador,
- Descrição,
- Fórmula/Retorno do resultado,
- Tipo do indicador qualitativo (ordinal ou nominal) ou quantitativo,
- Período de cálculo/análise,
- Período de atualização,
- Prioridade,
- Usabilidade,
- Utilidade na vertente de documentos, análises ou ações promocionais,
- Notas (em que são feitos breves comentários auxiliares),
- Existência, especificando se o indicador já existe na bases de dados da empresa.

Na vertente de análise dos resultados obtidos através do documento de requisitos, foram considerados critérios de seleção definidos em conjunto com a empresa para que o número de variáveis diminuísse. Smith (Smith, 2011) afirma que embora definida a prioridade registada dos requisitos no momento da entrevista, posteriormente, na análise destes, o critério de prioridade pode alterar-se devido a outros critérios levantados. No caso do projeto LAURIUS o mesmo aconteceu pois na seleção dos indicadores finais, para além da prioridade, foram tidos em consideração critérios como, necessidade comum de diferentes equipas e a usabilidade (Diária/Semanal). Adicionalmente, Smith (Smith, 2011) refere que outros componentes devem ser considerados para cada requisito, como por exemplo, quais os utilizadores que serão afetados ou qual o custo de implementação. Aplicado a este projeto, foi considerado outro componente mais adequado, o indicador de existência no sistema. Este indicador pode ser dividido em cinco patamares:

- existe,
- não existe e pode ser calculado,
- não existe mas entretanto foi avaliado, tornando assim possível a sua implementação,
- não existe e está em fase de avaliação de critérios e termos internos da empresa
- e não existe, considerado um requisito ‘fora da caixa’.

Estes critérios de seleção têm a vantagem de apoiar a garantia de que o produto final encontra-se o mais perto possível da necessidade comum das equipas envolvidas.

Fase 3: Exploração das ferramentas, abordagens e metodologias

Nesta fase, assim como na fase de Implementação, foram realizadas pesquisas sobre quais as ferramentas, abordagens e metodologias seriam apropriadas na criação da ferramenta. Tendo em conta que os utilizadores das existentes análises de dados e outros componentes possuem conhecimento numa ferramenta e numa linguagem utilizadas diariamente, foram essas as escolhidas para o desenvolvimento do projeto. SAS e SQL são as ferramentas e linguagens escolhidas para facilitar o processo de adaptação à nova ferramenta. Na implementação dos indicadores da tabela, foram consideradas metodologias em conformidade com critérios de eficácia, utilidade e organização.

Fase 4: Implementação da ferramenta Laurius

Como anteriormente referido neste documento, a metodologia ponderada para a Implementação da tabela foi a fragmentação/particionamento de dados analisada por Coronel e Morris (Coronel & Morris, 2016). Os autores exploram três tipos possíveis de particionamento de tabelas Horizontal, Vertical e Mixed. A Horizontal tem como vantagem dividir a tabela original em várias tabelas mantendo o número de colunas mas com linhas únicas. A Vertical, ao contrário da Horizontal, tem como vantagem manter o número de linhas garantindo a singularidade das colunas nas várias subtabelas. A Mixed como o próprio nome indica, combina as outras duas fragmentações dividindo a tabela original em várias linhas com colunas diferentes para cada uma.

Para o projeto LAURIUS, as grandes vantagens da metodologia de fragmentação/particionamento são o aumento do desempenho, o aumento do processamento do grande volume de dados, a redução do tempo de execução necessário para as consultas de dados e a simplificação em futura manutenção. Todavia, a integridade e nível de performance podem ser desvantagens a combater. Contudo, o particionamento Vertical é o mais adequado na implementação da estrutura de dados pois é necessário manter o número de linhas

associadas a todos os clientes do programa de fidelização, com os indicadores calculados em colunas únicas para todas as subtabelas. Esta solução permite que a construção do cálculo dos indicadores requisitados seja mais ágil, agregando assim várias pequenas tabelas para a criação da macro tabela LAURIUS.

4.2 Método seguido no projeto

Posteriormente à pesquisa das várias metodologias encontradas e tomada a decisão das mais adequadas ao projeto LAURIUS, esta secção detalhará as metodologias levadas a cabo descrevendo os diferentes passos e procedimentos em cada fase.

4.2.1 Definição do problema

Na primeira fase de Definição do problema, estabelecido para dois meses, dividiu-se em duas fases: a de aprendizagem do modelo de dados do programa de fidelização e a aprendizagem do modelo de negócio.

Inicialmente, em conjunto com a empresa, foram reunidos os conceitos que definiram o núcleo do desafio, estabelecendo assim o intuito da estrutura de dados. Foram revelados os objetivos da empresa, onde o principal objetivo é a grande utilidade que a macro tabela teria, determinados prazos flexíveis de outros objetivos e demonstrado que a base da estrutura seria a agregação de todos os requisitos das equipas gestoras do programa de fidelização.

A aprendizagem do modelo de dados definiu-se como um conhecimento autónomo e de exploração visto existirem restrições de acesso ao modelo original do sistema de informação atualmente implementado na empresa. Foram então estudados documentos, tabelas e ligações entre estas sobre as bases de dados analítica e operacional, existentes. Com o apoio dos elementos das equipas analíticas e gestoras dos dados armazenados, foi possível perceber determinadas associações e regras que deveriam ser respeitadas e utilizadas na implementação da estrutura de dados. De modo a organizar a informação adquirida foi desenhado, dentro da fase de Implementação mais à frente apresentada, um modelo de dados com as tabelas e associações entre estas, coerente com as ligações e regras existentes.

Na aprendizagem do modelo de negócio, as equipas gestoras do programa de fidelização foram envolvidas numa metodologia de observação no seu meio de trabalho, de forma a apresentarem individualmente a utilidade que os dados armazenados tinham. Análises de dados, análises de ações promocionais, definição de segmentações e relatórios são algumas das tarefas diárias destas equipas. Os objetivos destes outputs vão de encontro aos valores da empresa e da missão do programa de fidelização criado, que procura de uma forma simples, mas eficaz, manter o foco nas famílias abrangidas, compreendendo e satisfazendo assim os seus clientes. Perceber e questionar determinados conceitos, métricas calculadas e o que cada elemento necessitava, a nível de dados para executar as tarefas diárias na crescente satisfação do cliente, tornou-se essencial para a compreensão da utilidade da ferramenta a desenvolver.

4.2.2 Levantamento de Requisitos

A próxima fase passou por fazer o Levantamento de requisitos conjuntamente com as equipas utilizadoras da tabela desenvolvida. Através do método da entrevista semi-estruturada foram questionadas 10 equipas das várias áreas de decisão do programa de fidelização: Campaign Intelligence, LTP labs, Business Analytics, Advanced Analytics, Gestão de Parceiros, CCR (Customer Centric Retail) Categoria, CCR Operacional, CCR Cliente, Gestão Promocional e

Marketing Digital. A última equipa foi entrevistada no sentido de perceber que possíveis indicadores seriam relevantes armazenar futuramente a nível digital na estrutura de dados.

Durante o processo de entrevistas, foram registados todos os requisitos relevantes num documento de requisitos, de modo a compreender que necessidades existiam em cada equipa e procurando perceber quais os pontos importantes para o cálculo dos indicadores. A entrevista consistiu em perguntar ao utilizador quais os indicadores que calculava ou consultava regularmente, ou quais os indicadores que não existiam nas bases de dados e que seriam uma mais-valia caso existissem. Depois de registado o nome do indicador ou variável, o utilizador foi questionado quanto às restantes especificações (colunas) indicadas na tabela da documentação. Conforme apresentada no ponto anterior, a metodologia implementada nesta fase foi a documentação auxiliar criada. No Anexo C encontra-se uma imagem com alguns dos requisitos levantados, baseados numa tabela com cada um representado em cada linha relacionado com várias colunas. Apenas é apresentada parte do documento pois a sua importância reflete-se na primeira linha da tabela onde as colunas correspondem aos detalhes e especificações dos requisitos sendo estes o Número do requisito, Nome da equipa entrevistada, Nome do indicador, Descrição, Fórmula/Retorno, Período de cálculo/análise, Período de atualização, Prioridade, Usabilidade, Utilidade, Notas e Existência.

Seguidamente, analisando os resultados obtidos e documentados na tabela auxiliar, foram removidos requisitos repetidos, revelando uma necessidade comum de várias equipas, e foram elucidadas, junto destas, algumas dúvidas surgidas em relação a determinados detalhes, como o período de cálculo ou a usabilidade do requisito de forma a perceber o verdadeiro intuito deste.

Após a primeira filtragem de levantamento de requisitos, foi claro o principal problema existente. Muitas das necessidades encontradas já existiam calculadas mas eram tratadas de uma forma individual. A necessidade de calcular um indicador sempre que pedido, afetava as tomadas de decisão em relação à satisfação e comunicação com o cliente. Através do levantamento de requisitos, foi clara a necessidade da construção de uma ferramenta de indicadores atualizados, sustentando deste modo determinadas medidas ou providências. Assim, a análise das necessidades resultou em 198 indicadores (ver Anexo B) embora apenas alguns tenham sido selecionados, respeitando os critérios definidos de necessidade comum de diferentes equipas, prioridade e usabilidade (diária/semanal). Estes critérios, estabelecidos e apresentados à empresa na fase de seleção das variáveis, tiveram como base um dos grandes objetivos desta, a utilidade dos indicadores para as equipas de gestão de dados do cartão de fidelização. Para que a estrutura de dados criada fosse um projeto de sucesso, esta teria de ser útil para os seus utilizadores, cujos requisitos levantados são basicamente necessidades frequentes. Desta forma, para a ferramenta de dados foram selecionadas um total de oitenta e três variáveis/indicadores finais para implementação, apresentados detalhadamente no Anexo D. Tendo em conta as condições de implementação apresentadas no capítulo anterior e o período inicialmente definido para a realização do projeto, foi determinado que nem todos os oitenta e três indicadores seriam desenvolvidos, indicando para a empresa apenas o intuito e vantagens de cálculo de alguns deles. Como referido previamente, por motivos de restrição, determinados indicadores encontram-se com informação oculta.

Agrupando os indicadores, estes dividem-se em seis grupos de objetivos diferentes. Segundo os critérios de existência na base de dados da empresa apresentados anteriormente, nas figuras seguintes: a cor amarela indica que o indicador existe na base de dados da empresa; a cor laranja indica que o indicador não existe mas pode ser calculado; a cor verde indica que

entretanto foi avaliado, tornando assim possível a sua implementação; a cor violeta indica que o indicador não existe e está em fase de avaliação de critérios e termos internos da empresa e por fim a cor rosa claro o indicador não existe, considerado um requisito ‘fora da caixa’.

Relativamente a critérios de implementação, se o nome do indicador se apresentar com um fundo com cor azul, este encontra-se calculado na tabela LAURIUS, caso contrário não.

O grupo das Segmentações integra vinte e seis indicadores apresentados na Figura 2, com a finalidade da empresa segmentar de várias formas o cliente alvo. As opções do cliente como a hora preferencial de compra ou os tipos de produtos comprados são exemplos de segmentações relacionadas com o seu comportamento, que se encontram registadas por indicadores deste grupo.

Segmentações	
1	Segmentação 1
2	Segmentação 1 parceiros A,C,D U12m
3	Segmentação 3
4	Segmentação 4
5	Segmentação 5
6	Segmentação 6 parceiro A
7	Segmentação 4 top produtos
8	Segmentação 8
9	Segmentação 9
78	Segmentação 78
57	Filhos parceiro K
74	Gasto parceiro A
75	Valor cliente
29	Cliente brand
76	Cliente promoção scoring
71	Missão de compra
28	Balanço
16	Parceiro A preferencial
13	Loja preferencial
77	Contacto preferencial - channel scoring
14	Hora preferencial de compra
15	Dia de semana preferencial
80	Weather preferencial
19	Distância à loja preferencial
20	Distância à loja mais próxima
18	Categoria preferencial

Figura 2 – Grupo de indicadores de segmentação

Com o objetivo de reunir toda a informação de contacto com o cliente, foi criado o grupo de Contacto com seis indicadores ilustrado na Figura 3. Informação como a morada, *e-mail* ou número de telefone pode ser consultada mediante estes indicadores.

Contacto	
73	Mailable
10	Mobilable
11	Emailable
79	Appable
65	Correspondência devolvida
12	Nº de SMS enviados

Figura 3 - Grupo de indicadores de contacto com o cliente

A Figura 4 retrata os indicadores agrupados relacionados com a informação transaccional do cliente. Estes indicadores confirmam valores como as vendas ou os descontos das transações registadas.

<i>Informação transaccional</i>	
60	Transacções U12M
61	Transacções Umes
72	Devoluções
42	Vendas brutas U12M
43	Vendas brutas Umes
44	Vendas Líquidas U12M
45	Vendas Líquidas Umes
46	VL na DC Alimentar
47	VL na DC Frescos
48	VL na DC Têxtil
49	VL na DC Bazar
50	VL na DC Casa
51	VL na DC Foods & Bakery
52	VL online U12M
53	VL online Umes
54	VB online U12M
55	VB online Umes
56	VL dieta
37	Desconto acumulado U12M
38	Desconto acumulado Umes
39	Desconto de preço obtido U12M
40	Desconto de preço obtido Umes
83	Desconto acumulado expirado
41	Taxa desconto
66	Recency
59	Nº de lojas frequentadas
67	Nº de produtos distintos parceiro A
22	Cliente ativo no parceiro
25	Exclusivo parceiro

Figura 4 – Grupo de Indicadores de informação transaccional

Agregando indicadores alusivos à informação e dados pessoais do cliente, o grupo de indicadores Cliente da Figura 5 apresenta variáveis como o Género, Data de nascimento, Código postal entre outros campos.

<i>Cliente</i>	
23	Cliente Zona
21	Código Postal
32	Estado da conta
63	Género
64	Data de nascimento
30	Conta P
31	Conta L
26	Cliente novo cartão de fidelização
17	Cartão novo programa fidelização preferencial

Figura 5 - Grupo de indicadores de informação do cliente

A Figura 6 apresenta o grupo de indicadores que detêm informação promocional das transações registadas pelo cliente.

<i>Informação promocional</i>	
33	Nº de rebates U12M
34	Nº de não rebates
35	Nº de rebates U12M !por estratégia
36	Rebates sms notificação

Figura 6 – Grupo de indicadores com informação promocional

O último grupo de indicadores, ilustrado pela Figura 7, agrupa indicadores com informação adicional de segmentações, do cliente ou transaccional.

<i>Outros</i>	
24	Promoção W
58	Plano Y parceiro H
27	Plano de parceria Utilities
62	Saldo disponível
68	Nº de reclamações
69	Self-check-out
70	Utilizações Online
82	Clicks do e-mail
81	Nuvem clicks nos folhetos

Figura 7 – Grupo de indicadores com informação adicional

Respeitando os termos de confidencialidade de informação no presente documento e definido como um dos objetivos para a empresa e de investigação, alguns dos nomes de indicadores bem como dos parceiros desta, foram substituídos por variáveis, como por exemplo letras, e alguns dos resultados da tabela final foram alterados ou multiplicados por um valor aleatório. Deste modo foi preservada a segurança dos dados da empresa.

Do total de indicadores apresentados, foi decidido à priori que os últimos 10 da tabela do Anexo D não seriam implementados por se integrarem em dois grupos de não existência no sistema de informação atual, anteriormente mencionados: não existe mas está em fase de avaliação de critérios internos e não existe porque é um indicador 'fora da caixa'. Como previamente referido, nas figuras anteriores dos grupos de indicadores é possível verificar estes 10 indicadores com o fundo do respetivo número a cor violeta e rosa.

4.2.3 Exploração de ferramentas, abordagens e metodologias

A fase seguinte consistiu em explorar as ferramentas, abordagens e metodologias para o desenvolvimento da estrutura de dados. Nesta fase foi respeitado o foco nos utilizadores, considerando as ferramentas e linguagens usadas diariamente por estes. Durante a primeira fase de conhecimento do modelo de negócio e do modelo de dados, foi possível perceber as linguagens, as bases de dados e o *software* usado. Deste modo, para que a estrutura de dados desenvolvida tivesse um período de adaptação curto e rápido por parte das equipas, foram escolhidas a linguagem SQL, o *software* SAS e o sistema de gestão de base de dados Oracle Developer, todos eles instalados e utilizados pela empresa. Estas abordagens permitiram implementar a macro tabela de forma simples visto ser uma linguagem padrão utilizada em alguns sistemas de informação e não apresentar nenhum custo adicionado à empresa.

Com o objetivo de conhecer ainda melhor o sistema de informação atual da empresa, foram realizados alguns exercícios em SQL. Para além de aprofundar e relembrar conhecimentos desta linguagem, estes foram importantes na compreensão do modelo de dados existente porém simultaneamente restrito. Como conclusão, foi comprovado e decidido que era esta a linguagem mais indicada para a construção da ferramenta a desenvolver.

Para além da pesquisa sobre a parte prática da ferramenta, foi também considerada a parte teórica. A documentação que reflete a organização das tabelas criadas foi um dos conceitos importantes pesquisados e implementados na fase seguinte do projeto. Regras, associações e tabelas teriam de ser coerentes e simples para apoiar a compreensão dos utilizadores.

4.2.4 Implementação

Segue-se a fase de Implementação que se dividiu em quatro etapas: Extração de dados, Transformação de dados, Modelação de informação e Automatização. Determinadas as tarefas incorporadas nestas fases, estas foram executadas em paralelo visto se complementarem. Assim, foi estabelecido, como primeiro passo da extração de dados, a base de clientes em que os indicadores iriam atuar. A base de clientes para a realização de testes iniciais foi extraída e mantida numa tabela, para que o processamento dos cálculos dos indicadores fosse executado para os mesmos clientes, reduzindo também o tempo de execução necessário. Os critérios de filtro dos clientes ativos no programa de fidelização aplicados ponderaram todos os clientes que utilizam ou que se encontram registados apenas no cartão do programa de fidelização e que efetuaram compras no período máximo de cálculo dos indicadores. Os dados utilizados são referentes aos do cartão do programa de fidelização, pois confirmam clientes que utilizaram o cartão no acto da compra, o que não abrange os clientes que não usaram o cartão não existindo assim registo possível destes. Esta pesquisa e construção da tabela base resultou num total de aproximadamente três milhões de clientes, em que o período de compras aplicado foi de doze meses.

A validação dos dados foi um processo constante ao longo da implementação, garantindo a segurança, coesão e evitando repetições destes.

Definida a base de clientes e seguindo a metodologia escolhida sobre o particionamento das tabelas, foram determinadas as primeiras variáveis para cálculo. A construção dos indicadores na fase de transformação de informação teve sempre em conta os seguintes fatores: a existência do indicador, a reutilização de código e a complexidade a nível de criação e não no tempo de resposta. Para que este projeto fosse enriquecedor de conhecimento e valor, foram implementados, a nível de código, indicadores de três grupos de existência: (i) existente; (ii) não existente e há a reutilização de código; e (iii) não existente mas entretanto foi avaliado sendo possível o seu cálculo.

As variáveis não existentes ‘fora da caixa’, de acordo com a coluna de Prioridade do documento de requisitos, e as não existentes em fase de avaliação de critérios internos da empresa, não foram desenvolvidas. Para estas foi especificada uma descrição para futuras ações. Por exemplo, a variável número oitenta (*Weather* preferencial), da tabela do Anexo D, foi um requisito nunca antes pensado mas que seria interessante no futuro.

Assim, dando início ao desenvolvimento das variáveis, foram consultadas as especificações de existência e do tipo de retorno para cada requisito no documento de requisitos.

Foram criadas tabelas temporárias para que estas não ocupassem espaço desnecessário, disponibilizando assim espaço na memória do sistema e não afetando os processos em

paralelo das equipas. Para cada requisito foram programadas as consultas e os cálculos necessários com base nos *Schemas* e *Views* existentes na base de dados analítica. Uma *View* é uma pesquisa que reúne um conjunto de resultados armazenados numa *query*, disponibilizada depois aos utilizadores com acesso ao sistema para uso imediato. Um *Schema* é uma estrutura de armazenamento de dados, tabelas, consultas e outros componentes que sustentam a gestão da base de dados num ambiente restrito, pois admite poucas permissões de acesso e de escrita, embora não garanta suporte técnico. Assim, foi possível extrair, não só dados dos parceiros internos e externos do programa de fidelização e dados alimentados de forma automática pelo departamento com essa responsabilidade, como transações, clientes ou segmentações, extraíndo também código de algumas variáveis calculadas para a sua reutilização.

A nível de programação SQL, foram utilizados os comandos e a lógica adequada, de forma a agregar a tabela temporária à tabela base de clientes. Nas tabelas seguintes criadas, foram calculadas novas variáveis e depois utilizada a mesma metodologia de particionamento à tabela criada anteriormente, construindo assim a macro tabela LAURIUS. Todo o código escrito encontra-se indentado para fique mais organizado e perceptível aos utilizadores no futuro. Através da Figura 8 é possível verificar o esquema desta metodologia aplicada. A agregação permite que cada tabela mantenha os clientes singularizando as variáveis calculadas para cada um deles.

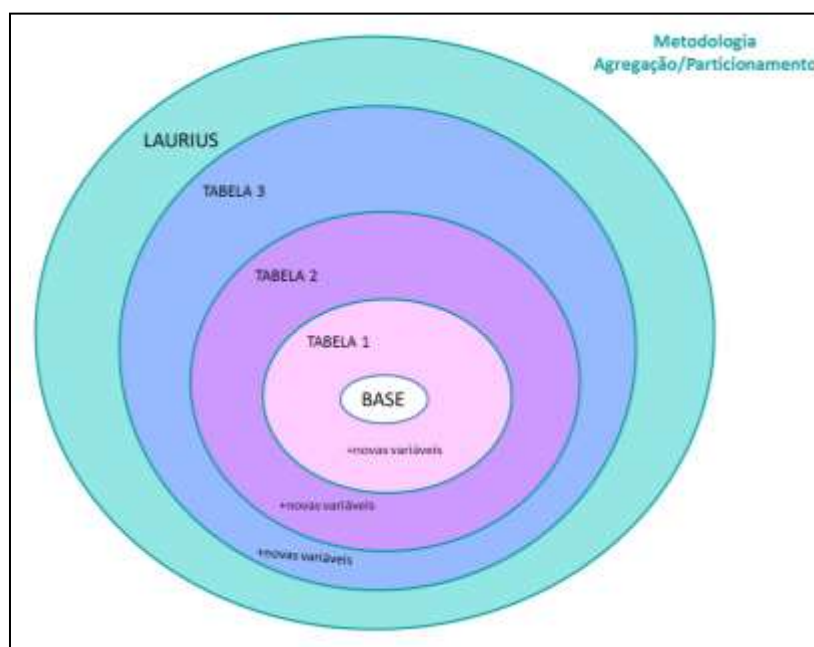


Figura 8 – Esquemática da metodologia de Agregação/Particionamento Vertical

A Figura 9 apresenta o desenho da metodologia da agregação implementada, com o exemplo da base de clientes agregada à tabela Particionamento 1 que por sua vez agrega-se à tabela Particionamento 2 e assim sucessivamente. A agregação tem como resultado a tabela macro LAURIUS.

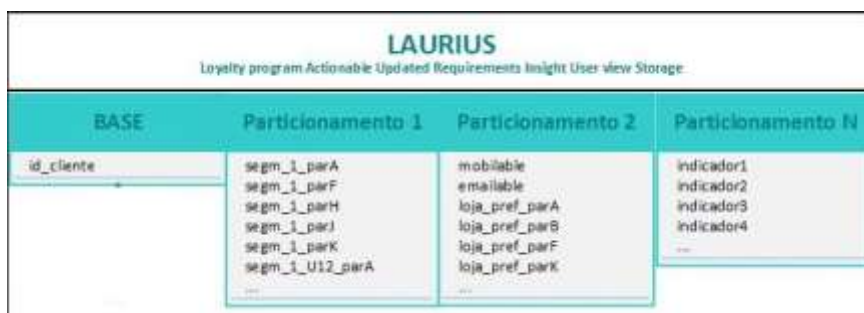


Figura 9 – Exemplo da metodologia aplicada na tabela LAURIUS

A metodologia aplicada foi a mais apropriada visto ser de rápida execução no cálculo individual das variáveis, bem como facilita a criação de futuras novas variáveis e a atualização dos vários indicadores ser efetuada de forma paralela.

Simultaneamente aos processos de cálculo e agregação de indicadores, foi desenvolvido um modelo de dados. Segundo a pesquisa realizada e com base no conhecimento adquirido na unidade curricular Sistemas de informação, o modelo de dados possui uma relação principal ternária e diversas relações entre as tabelas de particionamento e as tabelas temporárias de cálculo de indicadores ou as tabelas das bases de dados existentes. A relação ternária entre o cliente, o parceiro e o tempo de cálculo é essencial para o desenvolvimento das variáveis da tabela LAURIUS. A Figura 10 ilustra a relação ternária existente, em que cada indicador corresponde ao resultado de um dado cliente que faça compras num determinado parceiro e num período de tempo definido.

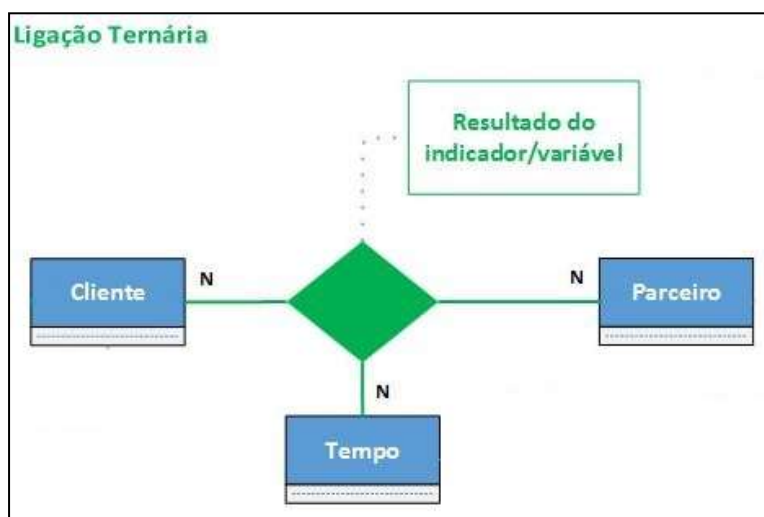


Figura 10 – Ligação ternária entre o Cliente, Parceiro e Tempo

Os Anexos E e F apresentam o modelo de dados geral e mais detalhado, respetivamente, tendo sido desenvolvidos na fase de modelação de informação. O modelo de dados geral ilustra a relação ternária existente, bem como a ligação com a tabela macro LAURIUS, onde a tabela da base de clientes é extraída pela tabela *customer* e as tabelas de particionamento são a agregação de vários resultados de indicadores calculados. No entanto, nem todos os indicadores utilizam esta ligação por serem estáticos e pessoais, não estando associados a nenhum parceiro, como o exemplo dos indicadores do género ou a data de nascimento.

O modelo de dados detalhado, apesar de se encontrar em anexo devido à sua dimensão, mostra a cor azul-escuro as tabelas dos *Schemas* e da *View* utilizadas para extrair os dados. Como uma amostra retirada do Anexo G, onde é exibido com maior detalhe todas as tabelas utilizadas para a extração referida, as Figuras 11 e 12 apresentam o *Schema* lab e a *View*, respetivamente.

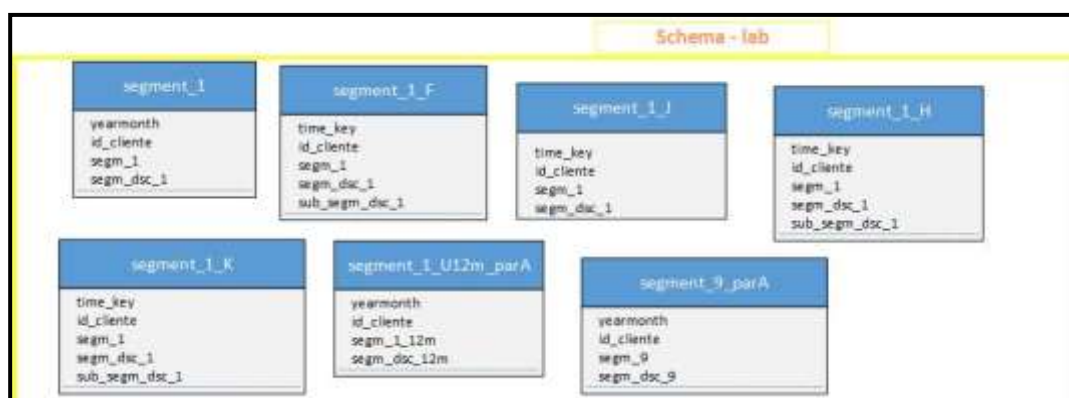


Figura 11 – Exemplo do *Schema* lab do Anexo G

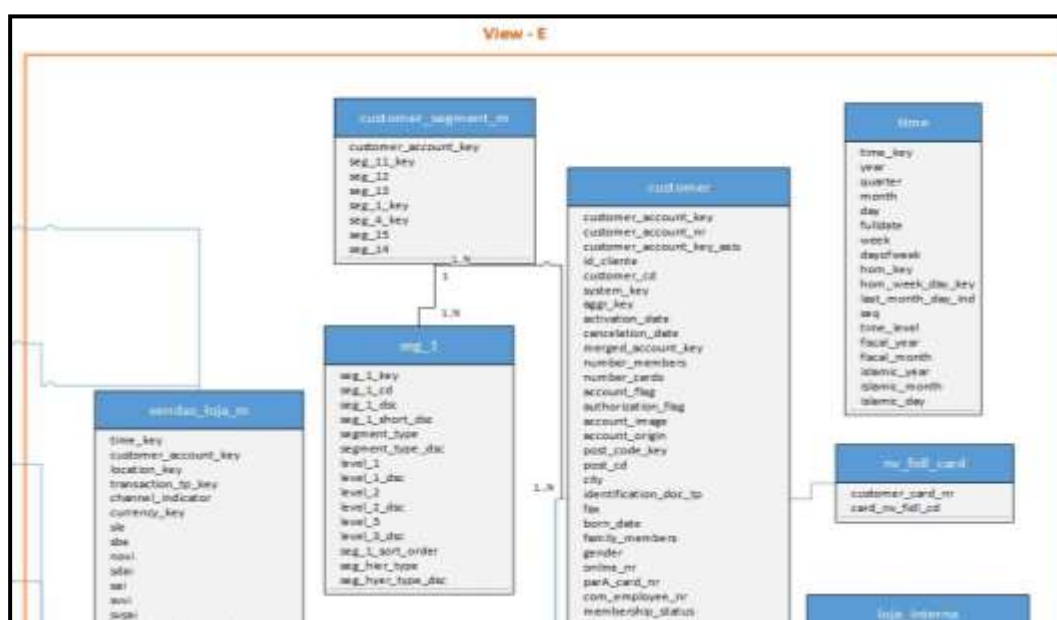


Figura 12 – Exemplo da *View* E do Anexo G

O Anexo H ilustra a verde as tabelas temporárias desenvolvidas para o cálculo individual dos indicadores e no Anexo I, a cor azul-claro, é possível observar o particionamento das tabelas para posterior agregação formando a macro tabela. As Figuras 13 e 14 exemplificam uma amostra dos anexos mencionados.

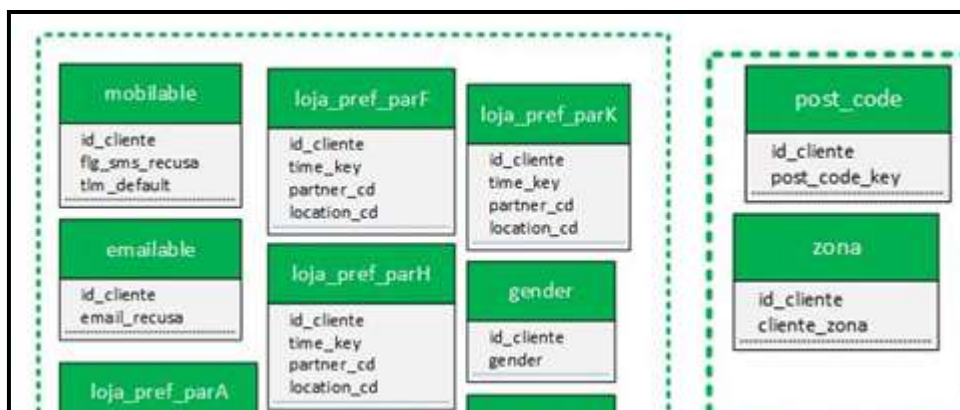


Figura 13 – Exemplo do Anexo H com algumas tabelas temporárias criadas

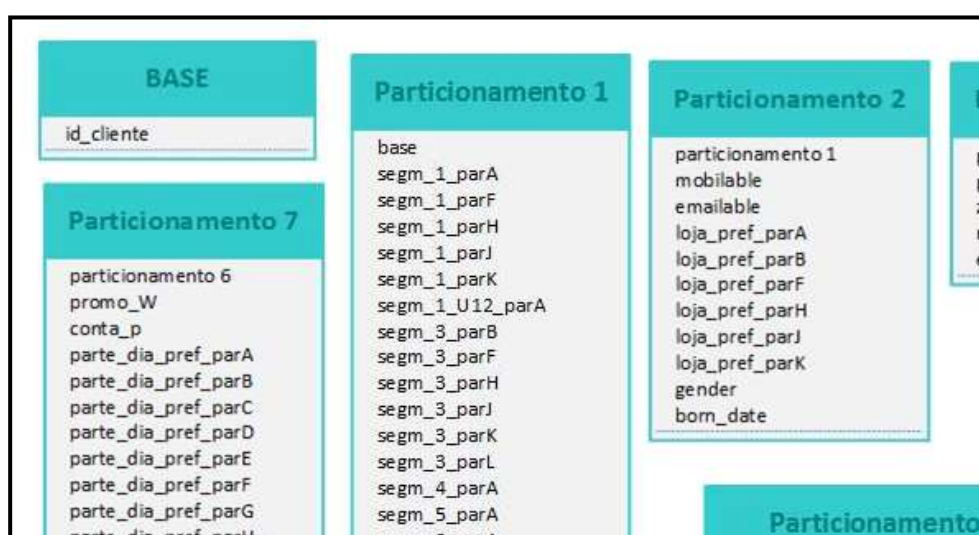


Figura 14 – Exemplo do Anexo I com algumas tabelas do particionamento para posterior agregação

Definida inicialmente como uma fase que englobava incerteza quanto à sua conclusão, a automatização dos dados não foi concluída devido ao curto período de desenvolvimento do projeto. No entanto, é um processo em curso com conclusão para breve.

Para além da validação de dados, ao longo da programação dos indicadores foi garantido que os campos das tabelas e o resultado do cálculo das variáveis retornassem sempre um valor que não fosse nulo. Algumas decisões em relação ao valor de retorno foram limitadas devido a políticas e metodologias aplicadas anteriormente, como por exemplo, o facto de o tipo de um campo ser do tipo número e não aceitar texto fez com que fosse tomada a decisão de substituir por um valor que não interferisse com o intuito e objetivo do indicador ou campo.

Como apoio para os utilizadores e de forma a facilitar as suas funções de análise de dados, foi adicionada ao final das colunas da tabela macro uma coluna com a data de atualização do cálculo dos indicadores, como ilustrado na Figura 15.

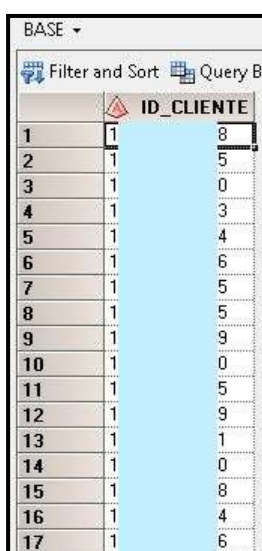
data_atualz
2017-05-03
2017-05-03
2017-05-03
2017-05-03

Figura 15 – Coluna com a data de atualização da tabela LAURIUS

5 Resultados

Neste capítulo é possível perceber os resultados das metodologias implementadas que foram descritas no capítulo anterior e a utilidade para a empresa de alguns dos indicadores implementados, mediante alguns exemplos. Os valores apresentados encontram-se adulterados de forma a preservar a informação confidencial da empresa.

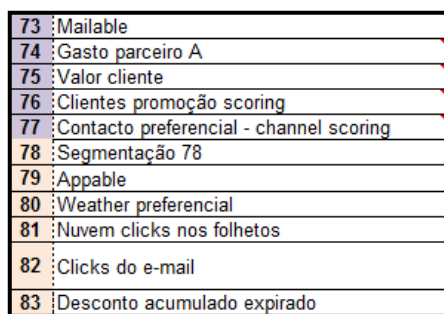
Na fase de implementação da macro tabela, foram definidos os critérios de filtragem dos clientes aderentes ao programa de fidelização. Resultando em aproximadamente três milhões de contas de clientes extraídas para a macro tabela LAURIUS, a Figura 16 apresenta uma amostra da tabela base de clientes ativos no programa de fidelização que realizaram compras nos últimos doze meses. Esta ilustra uma tabela com apenas uma coluna com os números de conta de todos os clientes conforme os filtros aplicados, ainda que, por motivos de confidencialidade, seja somente mostrado o primeiro e último, dígitos. Esta coluna é a base da macro tabela e é relevante para o objetivo de que cada linha da tabela represente um cliente detentor do programa de fidelização (User view).



	ID_CLIENTE
1	1 8
2	1 5
3	1 0
4	1 3
5	1 4
6	1 6
7	1 5
8	1 5
9	1 9
10	1 0
11	1 5
12	1 9
13	1 1
14	1 0
15	1 8
16	1 4
17	1 6

Figura 16 – Tabela criada com a base de clientes filtrados em SAS

Devido à sua dimensão, uma amostra da tabela LAURIUS encontra-se no Anexo I sendo possível afirmar que das oitenta e três variáveis, os onze indicadores da Figura 17 foram apenas definidos com texto, pois encontram-se ou em fase de avaliação de critérios internos da empresa ou são considerados ‘fora da caixa’ implicando uma maior pesquisa e análise do seu intuito e cálculo.



73	:Mailable
74	:Gasto parceiro A
75	:Valor cliente
76	:Clientes promoção scoring
77	:Contacto preferencial - channel scoring
78	:Segmentação 78
79	:Appable
80	:Weather preferencial
81	:Nuvem clicks nos folhetos
82	:Clicks do e-mail
83	:Desconto acumulado expirado

Figura 17 – Indicadores a roxo em fase de avaliação de critérios internos e a rosa indicadores ‘fora da caixa’

Destacando as 41 variáveis a azul da Figura 18 que se encontram implementadas na macro tabela, estas resultam em aproximadamente 200 colunas e cerca de três milhões linhas, onde estas podem variar conforme o número de clientes definido no período de compras da base de clientes.

Nº	Indicador	Nº	Indicador	Nº	Indicador
1	Segmentação 1	33	Nº de rebates U12M	65	Correspondência devolvida
2	Segmentação 1 parceiros A,C,D U12m	34	Nº de não rebates	66	Recency
3	Segmentação 3	35	Nº de rebates U12M, por estratégia	67	Nº de produtos distintos parceiro A
4	Segmentação 4	36	Rebates sms notificação	68	Nº de reclamações
5	Segmentação 5	37	Desconto acumulado U12M	69	Self-check-out
6	Segmentação 6 parceiro A	38	Desconto acumulado Umes	70	Utilizações Online
7	Segmentação 4 top produtos	39	Desconto de preço obtido U12M	71	Missão de compra
8	Segmentação 8	40	Desconto de preço obtido Umes	72	Devoluções
9	Segmentação 9	41	Taxa desconto	73	Mailable
10	Mobilable	42	Vendas brutas U12M	74	Gasto parceiro A
11	Emailable	43	Vendas brutas Umes	75	Valor cliente
12	Nº de SMS enviados	44	Vendas Líquidas U12M	76	Clientes promoção scoring
13	Loja preferencial	45	Vendas Líquidas Umes	77	Contacto preferencial - channel scoring
14	Hora preferencial de compra	46	VL na DC Alimentar	78	Segmentação 78
15	Dia de semana preferencial	47	VL na DC Frescos	79	Appable
16	Parceiro A preferencial	48	VL na DC Têxtil	80	Weather preferencial
17	Cartão novo programa fidelização preferencial	49	VL na DC Bazar	81	Nuvem clicks nos folhetos
18	Categoria preferencial	50	VL na DC Casa	82	Clicks do e-mail
19	Distância à loja preferencial	51	VL na DC Foods & Bakery	83	Desconto acumulado expirado
20	Distância à loja mais próxima	52	VL online U12M		
21	Código Postal	53	VL online Umes		
22	Cliente ativo no parceiro	54	VB online U12M		
23	Cliente Zona	55	VB online Umes		
24	Promocão W	56	VL dieta		
25	Exclusivo parceiro	57	Filhos parceiro K		
26	Cliente novo cartão de fidelização	58	Plano Y parceiro H		
27	Plano de parceria Utilities	59	Nº de lojas frequentadas		
28	Balanço	60	Transacções U12M		
29	Clientes brand	61	Transacções Umes		
30	Conta P	62	Saldo disponível		
31	Conta L	63	Género		
32	Estado da conta	64	Data de nascimento		

Figura 18 – Indicadores implementados a azul

Tendo em conta o grande volume de dados, o tempo de execução e atualização de dados gera uma média de oito horas. No entanto, uma melhoria futura está em dedicar um servidor por completo para que os tempos de execução e atualização necessária sejam otimizados e reduzidos.

Das dez equipas entrevistadas, a Campaign Intelligence, Advanced Analytics and Insight e Bussiness Analytics utilizam a maioria dos indicadores implementados para otimizar a criação de análises e implementar ações promocionais. Analisando a utilidade de alguns dos indicadores utilizados pela equipa Campaign Intelligence, os pedidos recebidos têm como objetivo o contacto direto com o cliente, como o envio de mensagens para o telemóvel, *e-mails* ou cartas por correio. Este procedimento respeita sempre um conjunto de regras definidas pela empresa de forma a gerar ações promocionais corretas e viáveis. Assim, exemplificando uma situação, para a realização do processo de envio de uma mensagem de telemóvel a divulgar uma ação promocional numa determinada loja de um parceiro, a equipa recorre às regras de envio especificadas. As regras chaves de contacto com o cliente respeitam sempre o contacto com o cliente em ações que lhe seja direcionado, respeita os limites de contacto estabelecidos, não permite o envio do mesmo contacto mais do que uma vez, respeita

a vontade do cliente no sentido de atender às reclamações e sugestões registadas por este e respeita ainda a sua privacidade, não ferindo a sua sensibilidade e precavendo a partilha e transporte de informação pessoal.

Aprovadas as regras chave do contacto direto com o cliente, é dado início ao processo de filtros e escrita da mensagem. Para a segmentação de clientes a quem a mensagem se destina, são utilizados indicadores como Mobilable, Cliente P e Cliente Zona. O primeiro indicador respeita a decisão do cliente ser contactado via mensagem de telemóvel e os outros dois filtram os clientes da categoria P e da Zona indicada onde se realiza a ação promocional. Através da consulta eficaz destes indicadores na tabela LAURIUS, todo o processo de envio de uma mensagem aumenta a sua eficácia de segmentação de clientes tendo em conta o objetivo da ação promocional, chegando assim rapidamente ao cliente a sua divulgação e informação sobre esta para que o cliente possa aderir ou usufruir de acordo com a periodicidade da ação promocional.

Por vezes, perceber se a mensagem de notificação enviada teve impacto na campanha ou nas vendas da loja, as equipas de Business Intelligence e Advanced Analytics and Insight criam análises com base nos dados do cartão do programa de fidelização. Estas têm como principal objetivo analisar o negócio de forma a perceber o impacto das várias campanhas e outras ações promocionais realizadas. Com a extração de toda a informação necessária, conforme o objetivo da análise, as equipas responsáveis pela ação promocional conseguem construir documentos na área de BI, como dashboards, utilizando metodologias e gráficos que permitem prever o comportamento do cliente, retirando depois *insights* na compreensão dos dados obtidos.

Exemplo 1:

Com o objetivo de perceber o impacto de cada canal nos clientes, as três equipas de CCR, Categoria, Cliente e Operacional, pretendem retirar conclusões da análise realizada ilustrada através da Figura 19.

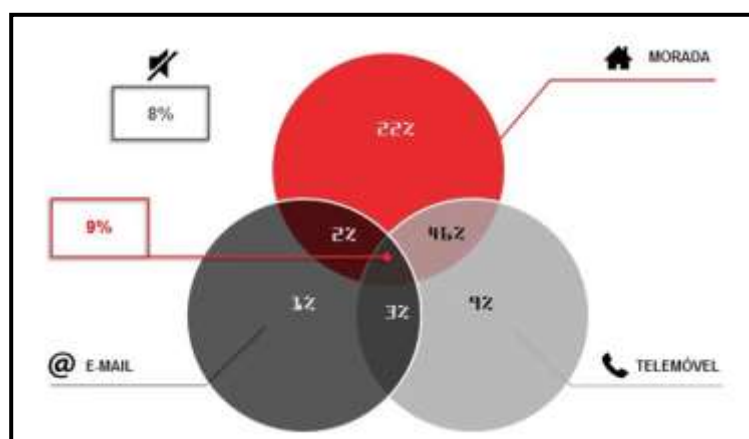


Figura 19 - Canais de contacto direto ao cliente

As decisões para estimular a utilização digital de outros canais é sustentada com base nos resultados obtidos. A Figura 19 apresenta um gráfico para cada canal de contacto direto com o cliente numa campanha temática. O canal que apresentou melhores resultados foi a carta enviada para a morada do cliente. No entanto, em conjunto com outro canal como o telemóvel ao enviar uma mensagem ao cliente, a percentagem é superior. Desta forma, em determinadas campanhas e em outras ações semelhantes, a escolha do canal ou combinações de canais de

contacto direto com o cliente é determinante para atingir resultados de sucesso tanto em vendas como em frequência ou em outros indicadores relacionados.

Exemplo 2:

Utilizando outros indicadores desenvolvidos na tabela LAURIUS, com o objetivo de desenvolver uma análise de vendas no último mês de clientes nas lojas do parceiro Q, as equipas responsáveis por aplicar os filtros necessários têm como base os indicadores Vendas brutas e a Hora preferencial de compra. A Figura 20 apresenta um gráfico desenvolvido onde é possível verificar esta informação filtrada da parte do dia preferencial de compra dos clientes das lojas do parceiro Q.

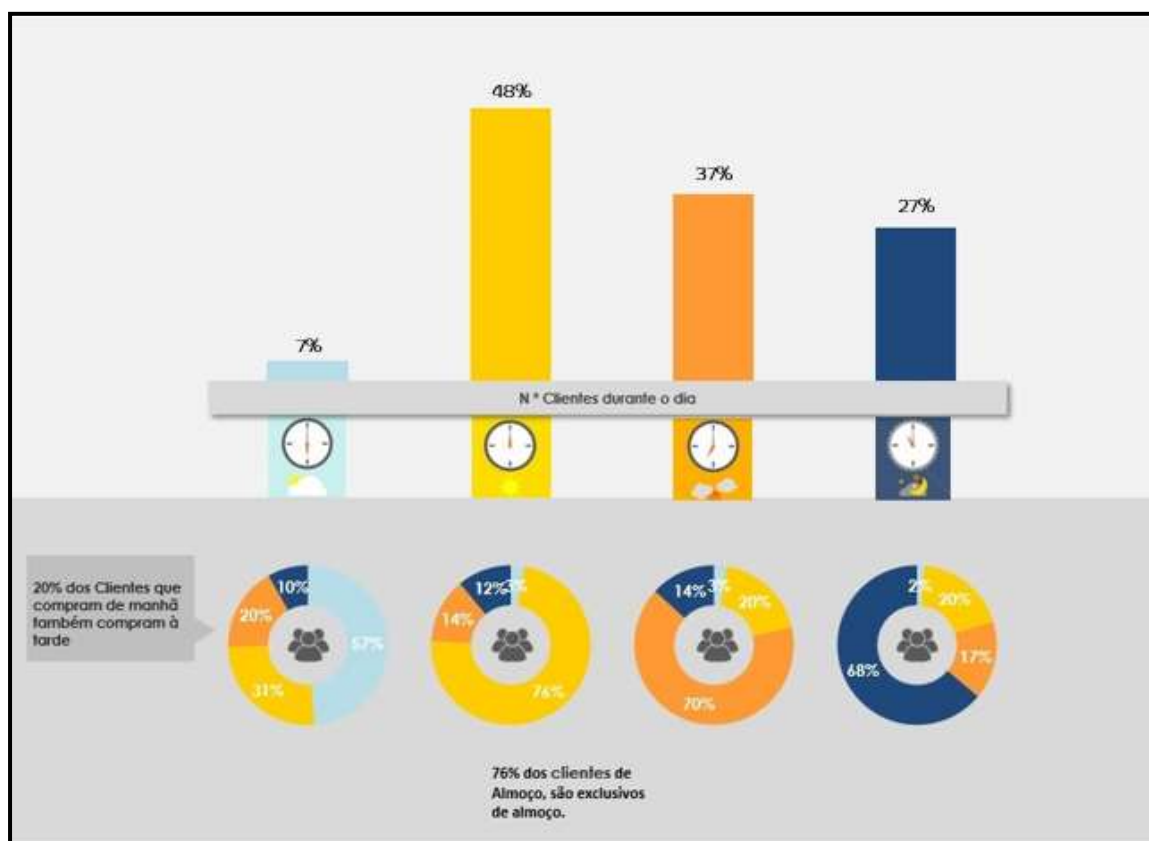


Figura 20 – Hora preferencial de compra no último mês nas lojas do parceiro Q

Sem a implementação dos respetivos indicadores da tabela LAURIUS como antes acontecia, este processo teria uma duração de desenvolvimento e análise de alguns dias tendo como consequências a perda de oportunidade de agir em conformidade com as acções da concorrência. Com a ferramenta implementada, os indicadores em questão encontram-se atualizados e podem ser consultados sempre que necessário.

Análises como esta permitem analisar o comportamento de compra nas quatro partes do dia definidas. Com base nesta análise, o conhecimento retirado possibilita às equipas de Gestão de Parceiros e as de CCR construir um plano de ações promocionais nas lojas em questão oferecendo aos clientes as melhores ofertas para os respectivos intervalos de compra preferenciais e procurar aumentar a frequência na cesta do cliente. Por exemplo, um cliente que esteja integrado na percentagem de clientes com compras no intervalo de hora preferencial do almoço, terá um plano de ofertas específico de modo a satisfazer os objetivos de compra do mesmo como adesão a campanhas promocionais ou descontos.

Exemplo 3:

O indicador número quatro, Segmentação 4, permite segmentar os clientes através do tipo de produtos comprados por estes nos últimos doze meses. A equipa de Advanced Analytics and Insight consegue desenvolver uma análise Ad-hoc de comparação entre o atual número e o índice de clientes com preferência da loja número dez do parceiro A. Os dados guardados nas bases de dados do homólogo em questão, permitem que sejam comparados de forma eficaz com os dados acionáveis e atualizados da macro tabela LAURIUS, filtrando assim os clientes dos indicadores segmentação 4 e loja preferencial. A Figura 21 ilustra um gráfico com o atual número e índice de clientes, por cada segmento da segmentação 4, relativo à loja número dez do parceiro A.



Figura 21 – Segmentação 4 dos clientes com loja preferencial 10 do parceiro A

Segundo esta análise, é possível destacar o segmento de clientes que mais frequenta esta loja, considerada preferencial pelos clientes filtrados, sustentando assim decisões de adaptação a este segmento do conceito, remodelação e evolução de loja e disposição de produtos. Deste modo, a loja pode moldar-se às preferências dos clientes do segmento destacado, no sentido da satisfação e comunicação com estes.

Exemplo 4:

Os dados referentes à mesma segmentação foram utilizados para uma análise de percentagem de vendas brutas dos clientes do parceiro A. A Figura 22 ilustra esta análise onde foi possível analisar dados segundo os cálculos dos indicadores da segmentação 4 e das vendas brutas para o parceiro A.

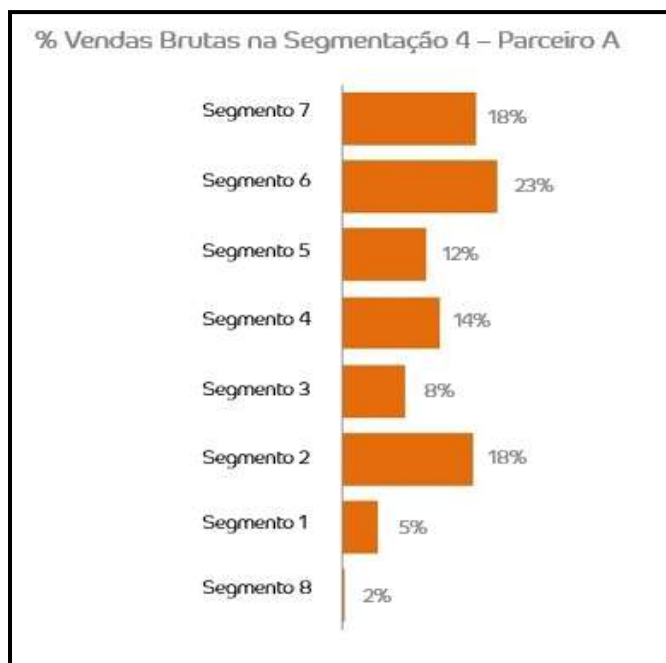


Figura 22 – Percentagem de vendas brutas na segmentação 4 dos clientes do parceiro A

A percentagem de vendas brutas é mais elevada nos clientes que pertencem ao segmento 6 existindo outros dois segmentos que se destacam mas com uma percentagem inferior. As conclusões retiradas desta análise realizada tiveram como objetivos refletir sobre quais os segmentos de clientes, dentro da segmentação 4, com maior volume de vendas durante uma campanha de venda de artigos temáticos, em relação aos resultados esperados nas estratégias implementadas.

Exemplo 5:

A Figura 23 cruza os resultados dos indicadores Género e Vendas líquidas na DC Casa. A análise pretende diferenciar os clientes no seu género, confirmando qual o que possui maior percentagem de vendas de artigos na direção comercial Casa do parceiro A.

Mediante os resultados obtidos é possível perceber alguns pontos de negócio importantes como: se os custos das ações tomadas é prejudicial nas vendas, se existem mais clientes angariados do que o esperado inicialmente e quais os pontos a melhorar na estratégia levada a cabo.

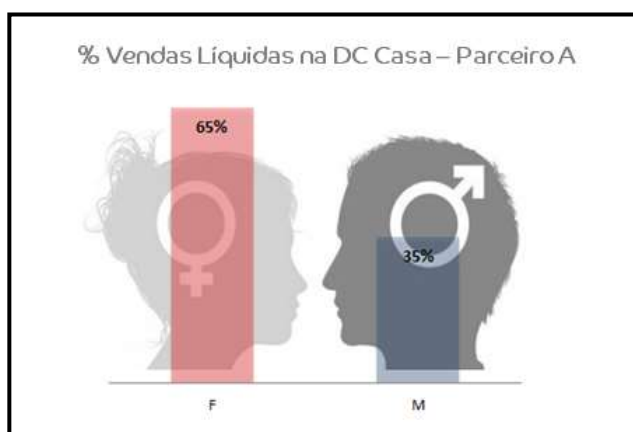


Figura 23 – Percentagem de vendas líquidas na DC Casa do parceiro A nos últimos 12 meses

6 Conclusões e perspectivas de trabalho futuro

Tal como foi referido ao longo do documento, no sector do retalho alimentar as margens de exigência do consumidor estão em constante crescimento, tornando valioso cada minuto de tempo “investido” na gestão de informação sobre o seu comportamento nas escolhas de compra, bem como na satisfação e contacto com o mesmo. Uma parte considerável desse investimento está associada às várias áreas de decisão da empresa como o departamento de marketing, programa de fidelização, área digital, análise comercial e operacional. A determinação de acesso pontual e de informação comportamental atualizada dos clientes utentes do programa de fidelização, implicou a identificação das necessidades e indicadores utilizados diariamente. Esta informação recolhida é essencialmente necessária para dar apoio na otimização da criação de análises e da implementação de ações promocionais das várias equipas envolvidas.

Num ambiente com um grande volume de dados e com diversos cálculos essenciais, a determinação de indicadores sempre que necessário, não permitia atuar de forma eficaz em conformidade com a satisfação do cliente e com o mesmo nível da concorrência. Desta forma, o objetivo do projeto para a organização passava por criar uma estrutura de Big Data atualizada *on-time* que servisse ao máximo as necessidades das diferentes equipas gestoras dos dados dos clientes detentores do cartão do programa de fidelização da empresa.

A ferramenta LAURIUS, como o nome indica, reúne todos os objetivos definidos inicialmente pela empresa. Assim, esta ferramenta proporciona uma melhor perspetiva para os utilizadores e contém todos os dados acionáveis e atualizados referentes ao programa de fidelização, todos os indicadores com base nas necessidades das equipas gestoras do mesmo programa de fidelização, possibilita a criação de análises no sentido de compreender melhor o comportamento do cliente, respeita uma visão cliente e viabiliza o armazenamento de todos os dados e indicadores necessários numa única estrutura.

Atendendo à metodologia utilizada no levantamento de requisitos, foram reunidas as necessidades das diversas equipas, como indicadores ou variáveis utilizados diariamente ou semanalmente, cumprindo deste modo o objetivo de compreensão clara e precisa de requisitos. Deste modo, foi garantido na triagem dos mesmos, os objetivos de eliminação de repetidos necessidade comum de diferentes equipas, prioridade e usabilidade destes.

Obedecendo às quatro particularidades apresentadas por Wedel e Kannan (Wedel & Kannan, 2016), a investigação caracterizou-se pelo volume de dados (Big Data) utilizado, implicando uma gestão assertiva e eficaz no desenvolvimento dos indicadores, pela velocidade de execução e desempenho, que podem ser futuramente otimizados quando a macro tabela estiver à responsabilidade de um servidor dedicado para o efeito, pela variedade e pela veracidade de indicadores e informação resultante. Além disso, todos os dados e informação utilizada cumpriram os critérios de coerência, segurança e confidencialidade estabelecidos.

A metodologia utilizada na construção da tabela LAURIUS deve ser repetida em processos e ferramentas semelhantes. Esta é organizada, metódica e focada no grande objetivo inicial de utilidade nos conceitos de Big Data e Customer Behavior. O particionamento revelou-se vantajoso na simplificação do processo de futura adição de novos campos assim como, em termos de resolução de problemas de deteção de erros ou falhas, é um procedimento simples e mais fácil na pesquisa e correção destes, e ainda aumenta o desempenho e funcionamento, melhorando o tempo de execução na consulta de dados exigida. O exemplo do envio de

mensagens para o telemóvel do cliente no processo de contacto com o cliente apresentado no capítulo dos resultados, revelou que a estrutura de dados desenvolvida viabiliza o acesso múltiplo e ágil à visualização de qualquer indicador necessário, através de pesquisas realizadas.

Paralelamente, e definido como um dos objetivos de investigação, foi desenvolvido um modelo de dados de apoio à macro tabela com intuito de tornar coerente e inequívoca a compreensão das relações entre as tabelas desenvolvidas e das bases de dados existentes.

Sendo uma empresa em constante evolução, há uma maior preocupação na valorização e satisfação do cliente, dando-lhe acesso a promoções ou descontos, otimizando a criação de análises de informação e sustentando as decisões na implementação de ações promocionais. Por estas razões, há melhorias que podem ser estudadas e implementadas na estrutura de dados desenvolvida. O código programado pode ser otimizado de modo a reduzir o tempo de execução. Outra melhoria a investigar seria a criação de um plano de implementação em Hadoop para ficar disponível e acessível para todas as equipas. A ferramenta Hadoop é uma plataforma de *software* em linguagem Java, de computação distribuída, com a vantagem de gerir clusters e processamento de grandes volumes de dados de forma rápida e flexível tendo um baixo custo para a empresa.

O desenvolvimento de uma interface poderá ser uma medida futuramente implementada de forma a melhorar a consulta de dados entre as equipas, aperfeiçoando e simplificando o processo de apresentação de resultados.

Uma melhoria em curso é a criação de um plano de atualização de dados mensal. No entanto, para alguns dos indicadores prioritários será realizada uma atualização diária ou semanal.

Concluindo, é possível afirmar que os resultados finais obtidos são muito satisfatórios, contribuindo assim positivamente para as necessidades da empresa. Espera-se que com a macro tabela LAURIUS a empresa seja capaz de tomar decisões sustentadas quanto à satisfação e contacto com o cliente, bem como realizar de forma ágil, análises versáteis e detalhadas sobre a informação obtida.

Referências

- Alter, S. (2002). The work system method for understanding information systems and information systems research.
- Caldeira, C. P. (2016). Introdução ao Modelo de Dados Relacional. Retrieved 08-05-2017, 2017, from <http://www.portalwebmarketing.com/Tecnologia/Introdu%C3%A7%C3%A3oaoModeloDeDadosRelacional/MDRVantagenseDesvantagensSistemasBaseDados/tabid/656/Default.aspx>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Q.*, 36(4), 1165-1188.
- Coronel, C., & Morris, S. (2016). *Database systems: design, implementation, & management*: Cengage Learning.
- Date, C. (1999). *An Introduction to Database Systems (Introduction to Database Systems)*. Pearson Education Inc.
- Dedić, N., & Stanier, C. (2016). *Measuring the Success of Changes to Existing Business Intelligence Solutions to Improve Business Intelligence Reporting*. Paper presented at the Research and Practical Issues of Enterprise Information Systems: 10th IFIP WG 8.9 Working Conference, CONFENIS 2016, Vienna, Austria, December 13–14, 2016, Proceedings 10.
- Evaluation Research Team, E. B. (2008, December). No.16 Data collection methods for program evaluation: Observation. Retrieved 03-05-2017, from <https://www.cdc.gov/healthyyouth/evaluation/pdf/brief16.pdf>
- Evaluation Research Team, E. B. (2009, January). No. 17 Data collection methods for program evaluation: Interviews. Retrieved 03-05-2017, 2017, from <https://www.cdc.gov/healthyyouth/evaluation/pdf/brief17.pdf>
- Golfarelli, M., Rizzi, S., & Cella, I. (2004). *Beyond data warehousing: what's next in business intelligence?* Paper presented at the Proceedings of the 7th ACM international workshop on Data warehousing and OLAP.
- Kocakoç, I. D., & Erdem, S. (2010). Business intelligence applications in retail business: OLAP, data mining & reporting services. *Journal of Information & Knowledge Management*, 9(02), 171-181.
- Kocakoç, I. D. a. E., Sabri (2010). Business Intelligence Applications in Retail Business: OLAP, Data Mining & Reporting Services. *Journal of Information & Knowledge Management*, Vol. 9, No. 2.
- Kotler, P., Armstrong, G., Saunders, J., & Wong, V. (1999). *Principles of Marketing - Second European Edition*.
- Kotler, P., Wong, V., Saunders, J., & Armstrong, G. (2005). *Principles of Marketing (4th European Edition)*: Edinburg: Prentice Education Limited, 2005.
- List, B., Bruckner, R. M., Machaczek, K., & Schiefer, J. (2002). *A comparison of data warehouse development methodologies case study of the process warehouse*. Paper presented at the International Conference on Database and Expert Systems Applications.
- Moura Borges, J., Galvão Dias, T., & Falcão e Cunha, J. (2015). *Modelação de Dados em UML - Uma Abordagem por Problemas*. Porto: FCA-Editora Informática.
- Pressman, R. S. (2005). *Software engineering: a practitioner's approach*: Palgrave Macmillan.
- Ranjan, J. (2009). Business intelligence: Concepts, components, techniques and benefits. *Journal of Theoretical and Applied Information Technology*, 9(1), 60-70.

- SAS, I. (2013). SAS Enterprise Guide. Retrieved 08-05-2017, from http://www.sas.com/en_us/software/enterprise-guide.html
- Smith, R. (2011). Writing a Requirements Document. Retrieved from CDL: Center J for W Distributed Learning: <http://www.cdl.edu/uploads/QdS>, 6.
- Stackowiak, R., Rayman, J., & Greenwald, R. (2007). Oracle Data Warehousing and Business Intelligence Solutions. Wiley Publishing, Inc, Indianapolis.
- Stoicescu, C. (2016). Big Data, the perfect instrument to study today's consumer behavior. *Database Systems Journal*, 6(3), 28-42.
- Wedel, M., & Kannan, P. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97-121.

Bibilografia

- Abuelyaman, E. S. (2008). An optimized scheme for vertical partitioning of a distributed database. *IJCSNS International Journal of Computer Science and Network Security*, 8(1), 310-316.
- Applebaum, W. (1951). Studying customer behavior in retail stores. *Journal of Marketing*, 16(2), 172-178.
- Batini, C., Lenzerini, M., & Navathe, S. B. (1986). A comparative analysis of methodologies for database schema integration. *ACM computing surveys (CSUR)*, 18(4), 323-364.
- Hacigumus, H., Iyer, B., & Mehrotra, S. (2002). Providing database as a service. Paper presented at the Data Engineering, 2002. Proceedings. 18th International Conference on.
- Harrington, T. J. (2016). Merges and Joins. *Coders' Corner*, Paper 109-125.
- Infopédia dicionários porto editora: <https://www.infopedia.pt/dicionarios/lingua-portuguesa/vender>. Acesso a 20 de Março de 2017.
- Maffezoli, P. (2009). *Micromotives and Macrobehavior*-Thomas Schelling. Humana.Mente.
- Moura Borges, J., Galvão Dias, T., & Falcão e Cunha, J. (2015). *Modelação de Dados em UML - Uma Abordagem por Problemas*. (pp. 192) Porto: FCA-Editora Informática.
- Navathe, S. B., & Ra, M. (1989). Vertical partitioning for database design: a graphical algorithm. *SIGMOD Rec.*, 18(2), 440-450. doi: 10.1145/66926.66966
- Negash, S. (2004). *Business intelligence*.
- Pistor, P., & Traunmueller, R. (1986). A database language for sets, lists and tables. *Information systems*, 11(4), 323-336.
- Rob, P., & Coronel, C. (2007). *Database Systems: Design, Implementation, and Management*: Cengage Learning.
- Slavakis, K., Giannakis, G. B., & Mateos, G. (2014). Modeling and optimization for big data analytics:(statistical) learning tools for our era of data deluge. *IEEE SIGNAL PROCESSING MAGAZINE*, 31(5), 18-31.
- Uncles, M. D., Dowling, G. R., & Hammond, K. (2003). Customer loyalty and customer loyalty programs. *Journal of Consumer Marketing*, 20(4), 294-316.
- Wang, R. Y., Kon, H. B., & Madnick, S. E. (1993). Data quality requirements analysis and modeling. Paper presented at the Data Engineering, 1993. Proceedings. Ninth International Conference on.
- Zakaria, I., Rahman, B. A., Othman, A. K., Yunus, N. A. M., Dzulkpli, M. R., & Osman, M. A. F. (2014). The relationship between loyalty program, customer satisfaction and customer loyalty in retail industry: A case study. *Procedia-Social and Behavioral Sciences*, 129, 23-30.

ANEXO A: Planeamento do Projeto LAURIUS

Planeamento Projeto LAURIUS				Fevereiro				Março				Abril				Maio				Junho				Julho				
Dia 1				6	13	20	27	6	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19	26	3	10	17	24
Dia 7				12	19	26	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18	25	2	9	16	23	30
Realização da Tese																												
Definição do problema																												
-Aprendizagem do modelo de dados do programa de fidelização.																												
-Aprendizagem do modelo de negócio																												
Levantamento de requisitos.																												
-Análise dos resultados (variáveis repetidas, se existem, em que tabelas existem)																												
-Seleção de variáveis (selecção das variáveis conforme os critérios definidos)																												
Exploração das ferramentas, abordagens e metodologias																												
Implementação SAS-SQL 83 Variáveis																												
- Extração de dados (definição da base de clientes)																												
- Transformação de dados (desenvolvimento do cálculo das variáveis, reutilização de código)																												
-Modelação de informação (desenvolvimento adaptação contínua do modelo de dados)																												
-Automatização MBM																												
Data limite para envio do título definitivo da dissertação 7 Junho																												
Entrega da versão provisória da dissertação para revisão 16 Junho																												
Correções na tese																												
Entrega da versão definitiva (antes da apresentação) em PDF 30 Junho																												
Preparação da apresentação																												
Apresentações públicas 17 a 21 de julho																												
Versão final em PDF num CD em caixa normal rígida + formulário de entrega da dissertação preenchido e assinado 28 Julho																												

ANEXO B: 198 Indicadores recolhidos na fase de Levantamento de requisitos

<i>Levantamento de Requisitos Macro-Behavior Customer MasterData</i>			
Nº	Equipa	Variável	Descrição
1	Campaign Intelligence	Segmentação 1	Segmentação 1 é calculada para os parceiros A, C, D, F, H, J e K tendo como objetivos compreender o impacto de novos formatos de loja em cada nível de valor e identificar potenciais clientes de alto valor que comprem noutros parceiros.
2	Campaign Intelligence	Segmentação 3	Segmentação 3 é classificada pela frequência do cliente de acordo com o parceiro.
3	Campaign Intelligence	Mobilable	Flag com dois possíveis retornos definindo se o cliente permite ou não ser contactado via telemóvel.
4	Campaign Intelligence	Emailable	Flag com dois possíveis retornos definindo se o cliente permite ou não ser contactado via email.
5	Campaign Intelligence	Mailable	Flag com dois possíveis retornos definindo se o cliente permite ou não ser contactado via correio. fase de avaliação de critérios internos do seu cálculo e intuito.
6	Campaign Intelligence	POSalable	Flag se o cliente utiliza POS

7	Campaign Intelligence + CCR Cliente	Appable	Quem usou a aplicação móvel do cartão de fidelização.
8	Campaign Intelligence + CCR Cliente	Tiravezable	Quem usou Tira-vez.
9	Campaign Intelligence	Cliente ativo no parceiro	Se fez compras nos últimos 12 meses (mesmo que seja uma compra de 10 cêntimos).
10	Campaign Intelligence +LTP	Cliente Zona (Ilhas, Espanha)	Saber se o cliente é das ilhas (Açores ou Madeira) ou Portugal Continental. Avaliar se possível zona fora de Portugal. Analisar morada e loja preferencial
11	Campaign Intelligence	Promoção W	Saber se adere à promoção W
12	Campaign Intelligence	Cliente P	Saber se é um cliente P
13	Campaign Intelligence	Cliente L	Saber se é cliente L
14	Campaign Intelligence	Género	Saber se é masculino ou feminino
15	Campaign Intelligence	Idade	Data de nascimento; *saber faixa etária

16	Campaign Intelligence	Agregado Familiar	Número de pessoas que vivem na mesma casa e com alguma relação de parentesco.
17	Campaign Intelligence	Código Postal	Código postal - post_cd .
18	Business Analytics	Código Postal	Código postal - post_code_key .
19	Campaign Intelligence	Nº de rebates nos últimos 12 meses	Distinguir rebates por Parceiro; Indicar qual o parceiro. Atenção ao canal de contacto.
20	Campaign Intelligence	Plano Y	Saber se cliente é ilegível ou se já usou Plano Y parceiro H.
21	Advanced Analytics & Insight	Utilizadores App	Frequência da utilização da app; Possível segmentação.
22	Advanced Analytics & Insight	Utilizadores Tira-vez	Frequência da utilização do tira-vez Possível segmentação.
23	Advanced Analytics & Insight	Nº de clientes F	Número de clientes F.

24	Business Analytics	Clientes exclusivos de cada parceiro	Valor de vendas por parceiro, caso apenas 1 (parceiro) esteja preenchido é exclusivo desse parceiro.
25	Business Analytics	Loja preferencial	Loja preferencial por parceiro (de acordo com o parceiro).
26	Business Analytics	Segmentação 4	Análise por tipologia de produtos nos últimos 12 meses para os parceiros A, C e D.
27	Business Analytics	Segmentação 5	Com base no comportamento de compra do cliente, para os parceiros A, C e D. Nome do segmento atribuído ao retorno do cálculo dividido em 3 patamares de afinidade a produtos de crianças.
28	Business Analytics	Cliente têxtil	Dividir têxtil infantil e adulto, por parceiro (J, K, A, C e D).
29	Business Analytics	Cliente por canal	Clientes que usaram mensagem de telemóvel, carta ou POS por parceiro.
30	Business Analytics	Cliente do novo cartão de fidelização	Flag retornando se é cliente do novo cartão de fidelização.
31	Business Analytics	Pagamento com novo cartão de fidelização	Pagamentos (vendas) por parceiro.
32	Business Analytics	Distância à loja preferencial	Loja preferencial definida (de acordo com o parceiro).
33	Business Analytics	Hora preferencial de compra	Por parceiro; Horas/parte do dia em que o cliente prefere fazer compras.

34	Business Analytics	Dia de semana preferencial	Por parceiro; Dia da semana que o cliente prefere fazer compras.
35	Business Analytics	Dia do mês preferencial	Por insígnia/parceiro; Dia do mês que o cliente prefere fazer compras.
36	Campaign Intelligence - LTP	Segmentação 6	Segmentação às lojas próprias VS lojas concorrência (distâncias).
37	Campaign Intelligence + LTP	Preferência cupão valor Continente	Preferência identificada pelo cliente através do call center.
38	Campaign Intelligence + LTP	Nº de SMS enviados	Contador de mensagens enviadas. Por parceiro;
39	Campaign Intelligence +LTP	Nº POS emitidos	Número de POS emitidos.
40	Campaign Intelligence + LTP	Estado da conta	Saber o estado da conta conforme os tipos existentes.
41	Campaign Intelligence+ LTP	Presença CRM	Saber se cliente está registrado no CRM.
42	Campaign Intelligence + LTP	Correspondência devolvida	Identificar moradas inválidas ou com falhas.
43	Campaign Intelligence + LTP	Cliente associou novo cartão de fidelização	Se o cliente associou cartão fidelização ao novo cartão de fidelização.
44	Business Analytics	Filhos parceiro K	Em concordância com os dados dos clientes do parceiro K, número de filhos do cliente.

45	Business Analytics	Weather preferencial	Weather na loja (em todas as lojas); Sol, chuva, temperatura min e vento; Por parceiro; Dificuldade ++
46	Business Analytics	Recency	Número exato de dias que o cliente não fez compras, por parceiro.
47	Business Analytics	Plano de parceria Utilities	Todos os clientes que estão habilitados a receber Plano de parceria Utilities.
48	Business Analytics	Distância loja mais próxima	Distância à loja mais próxima por parceiro. Georreferenciação.
49	Business Analytics	VL DC 1 Parceiro H	Vendas Líquidas Direcção Comercial 1 Parceiro H.
50	Business Analytics	VL DC 2 Parceiro H	VL DC 2 Parceiro H.
51	Business Analytics	VL DC3 Parceiro H	VL DC3 Parceiro H.
52	Business Analytics	Clientes V	Flag se comprou nos últimos 12 e não nos últimos 12 anteriores.
53	Business Analytics	Clientes D	Flag se comprou nos últimos 12 anteriores e não nos últimos 12.
54	Business Analytics	Clientes T	Flag se comprou nos últimos 12 e não nos últimos 12 anteriores.
55	Business Analytics	Categoria de entrada	Primeira compra dos clientes, por parceiro.

56	Business Analytics	Nº de lojas frequentadas	Número de lojas frequentadas, por parceiro.
57	Business Analytics	Nº de utilizações Online	Quantos já utilizaram o serviço on-line por parceiro (H, K, J, A, C e D).
58	Business Analytics	Self-check-out	Número de transações de self-check-out pelo número total de transações do cliente.
59	Advanced Analytics & Insight	Nº de parceiros	Número de parceiros por cliente.
60	Business Analytics	Cartão novo programa de fidelização Preferencial	Cartão associado ao cliente atual.
61	Business Analytics	Account activation date	Data de ativação do cartão.
62	Business Analytics	Account origin	Conta origem.
63	Business Analytics	Vendas brutas	Vendas brutas por parceiro e total do ecossistema = Valor original sem descontos. Últimos 12 meses.
64	Business Analytics	Vendas brutas	Vendas brutas por parceiro e total do ecossistema = Valor original sem descontos. Último mês.

65	Business Analytics	Vendas Líquidas	VL por parceiro e total do ecossistema = Vendas Brutas menos o Imposto sobre o Valor Acrescentado. Últimos 12 meses.
66	Business Analytics	Vendas Líquidas	VL por parceiro e total do ecossistema = Vendas Brutas menos o IVA. Último mês.
67	Business Analytics	Transações	Número de Transações por parceiro e total do ecossistema.
68	Business Analytics	Transações	Número de Transações por parceiro e total do ecossistema
69	Business Analytics	Quantidade	Quantidade por parceiro e total do ecossistema.
70	Business Analytics	Quantidade	QTY por parceiro e total do ecossistema.
71	Business Analytics	Cesta média em Vendas Brutas	Cesta média em VB por parceiro e total do ecossistema.
72	Business Analytics	Cesta média em Quantidade	Cesta média em QTY por parceiro e total do ecossistema.
73	Business Analytics	Gasto médio Vendas Brutas	Gasto médio em VB por parceiro e total do ecossistema.
74	Business Analytics	Desconto acumulado	Por parceiro e total do ecossistema.
75	Business Analytics	Desconto acumulado	Por parceiro e total do ecossistema.
76	Business Analytics	Pagamentos Saldo usado	Pagamentos com o saldo do cartão de fidelização por parceiro e total do ecossistema. Últimos 12 meses.

77	Business Analytics	Pagamentos Saldo usado	Pagamentos com o saldo do cartão de fidelização por parceiro e total do ecossistema. Último mês.
78	Business Analytics	VL na DC Alimentar	Vendas Brutas menos o IVA na DC- direção comercial Alimentar. Parceiros A, C e D.
79	Business Analytics	VL na DC Frescos	Vendas Brutas menos o IVA na DC- direção comercial Frescos Parceiros A, C e D.
80	Business Analytics	VL na DC Têxtil	Vendas Brutas menos o IVA na DC- direção comercial Têxtil Parceiros A, C e D.
81	Business Analytics	VL na DC Bazar	Vendas Brutas menos o IVA na DC- direção comercial Bazar Parceiros A, C e D.
82	Business Analytics	VL na DC Casa	Vendas Brutas menos o IVA na DC- direção comercial Casa Parceiros A, C e D.
83	Business Analytics	VL na DC Foods & Bakery	Vendas Brutas menos o IVA na DC- direção comercial Foods & Bakery. Parceiros A, C e D.
84	Business Analytics	VL na categoria Bacalhau	Vendas Brutas menos o IVA do Bacalhau.
85	Business Analytics	Peso QTY premium plano de parceria Utilities	Peso das QTY premium no total de QTY plano de parceria Utilities.
86	Business Analytics	Desconto de preço acumulado	Desconto no preço, por parceiro e total do ecossistema. Últimos 12 meses.
87	Business Analytics	Desconto de preço acumulado	Desconto no preço, por parceiro e total do ecossistema.

			Último mês.
88	Business Analytics	(Categoria preferencial) TOP 5 produtos afinidade	Categoria preferencial, por parceiro; Categoria com mais QTY ou preço mais alto na transação; Parceiros A, C e D.
89	Business Analytics	DTCCFR	Código da Freguesia- Distrito Concelho e Freguesia
90	Business Analytics	ID cliente look-a-like	Cliente igual, segundo os critérios em avaliação por parceiro ou campanha.
91	Business Analytics	Variação de VL por parceiro	Informação da variação de vendas líquidas por parceiro.
92	Business Analytics	Nº de mailings recebidos	Número de mailings recebidas, por parceiro.
93	Business Analytics	Nº de mailings que participou	Número de mailings que participou, por parceiro.
94	Business Analytics	Data da última adesão promoção FS	Data da última adesão promoção FS do cartão de fidelização.
95	Business Analytics	Nº de não rebates	Número de cupões passados em caixa mas não rebatidos.
96	Business Analytics	Vendas associadas ao valor de preço	Vendas Líquidas quando usado o desconto de preço.
97	Business Analytics	Vendas mailings	Vendas em mailings nos Parceiros A,C e D.

98	Business Analytics	Descontos em mailings	Número de descontos em carta Parceiros A,C e D.
99	Business Analytics	Vendas associadas a descontos	Vendas quando usado descontos.
100	Business Analytics	Saldo disponível	Saldo disponível no cartão de fidelização.
101	Business Analytics	Desconto atribuído em Plano P	Valor disponível com data limite.
102	Business Analytics	Desconto atribuído não utilizado em Plano P	Valor disponível com data limite que não foi utilizado.
103	Advanced Analytics & Insight	Nº de rebates nos últimos 12 meses, !por estratégia	Consoante as estratégias definidas.
104	Advanced Analytics & Insight	Nº de produtos distintos	Número de produtos diferentes Parceiros A,C e D.
105	Advanced Analytics & Insight	Nº de semanas em que o cliente fez compras	Número de semanas que o cliente fez compras Parceiros A,C e D.
106	Advanced Analytics & Insight	Nº de cupões de produtos não rebatidos	Contador de cupões não rebatidos;
107	Advanced Analytics & Insight	Distância à loja mais próxima ao concorrente	Distância entre o parceiro A e o concorrente mais próximo.
108	Advanced Analytics & Insight	Distância entre clientes	Distância entre clientes que aderentes ao cartão de fidelização.

109	Advanced Analytics & Insight	Dependente Plano Y parceiro H	Quais são os clientes que dependem desse Plano Y parceiro H.
110	Advanced Analytics & Insight	Cliente channel scoring	Meio de contacto preferencial do cliente: telefone, email,... Engagement por canal; Por tipo de oferta.
111	Advanced Analytics & Insight	Gasto parceiro A	Percentagem gasta em orçamento alimentar em Parceiro A.
112	Advanced Analytics & Insight	Data da última compra	Por parceiro.
113	Advanced Analytics & Insight	Balanço	Flag acionada para o cliente propenso ao abandono.
114	Advanced Analytics & Insight	Segmentação 10	Segmentação 10
115	Advanced Analytics & Insight	Segmentação 11	Sim ou não, e o valor associado a probabilidade a ser do segmento definido.
116	Advanced Analytics & Insight	Peso dos descontos totais nas compras do cliente	Peso dos descontos totais nas compras
117	Advanced Analytics & Insight	Top afinidades	Top de afinidade de produtos com base nos níveis definidos.
118	Advanced Analytics & Insight	VL on-line	Vendas líquidas de quanto cliente gastou on-line, últimos 12 meses.
119	Advanced Analytics & Insight	VL on-line	Vendas líquidas de quanto cliente gastou on-line, último mês.
120	Advanced Analytics & Insight	VB on-line	Vendas brutas de quanto cliente gastou on-line últimos 12 meses.

121	Advanced Analytics & Insight	VB on-line	Vendas brutas de quanto cliente gastou on-line último mês.
122	Advanced Analytics & Insight	Loja de conveniência	Loja de conveniência por cliente; Considerar QTY ou preço
123	Advanced Analytics & Insight	Loja de abastecimento	Loja de abastecimento por cliente; Considerar QTY ou preço.
124	Advanced Analytics & Insight	Valor cliente	Valor do cliente a longo prazo; Em fase de avaliação de critérios internos do seu cálculo e intuito.
125	Advanced Analytics & Insight	Nº de sugestões	Número de sugestões efetuadas.
126	Advanced Analytics & Insight	Nº de reclamações	Número de reclamações efetuadas.
127	Gestão de Parceiros	Filhos parceiro A	Número de filhos segundo segmentação 5 para parceiro A.
128	Gestão de Parceiros	Regresso aulas	Clientes que comprem artigos de criança, manuais escolares, livros juvenis ou apoio escolar; Clientes que comprem papelaria.
129	Gestão de Parceiros	Parceiro A preferencial	Clientes com loja preferencial parceiro A adjacente às lojas por parceiro.

130	Gestão de Parceiros	Lojas adjacentes	Proximidade de lojas entre parceiros Lojas por galeria ou concelho.
131	Campaign Intelligence	Contacto preferencial	Meio de contacto preferencial do cliente: telefone, email,...
132	Campaign Intelligence	Melhor oferta	Melhor oferta para o cliente.
133	Campaign Intelligence	Preferência do conceito de loja	Hipermercados, minimercados.
134	Campaign Intelligence	Sacos	Saber se o cliente compra sacos.
135	Campaign Intelligence	Cliente social	Saber se o cliente frequenta redes sociais.
136	Campaign Intelligence	Cliente afinidade marca	Se o cliente frequenta os eventos parceiros A, C e D.
137	Campaign Intelligence	Self shop	Se o cliente usa o serviço self scan.
138	Campaign Intelligence	Proximidade transportes públicos	Transporte público próximo da loja (metro, autocarro...).
139	Campaign Intelligence	Clicks do e-mail	Se o cliente clicou na hiperligação do e-mail enviado.
140	Campaign Intelligence	Senhas	Senhas da padaria ou peixaria ou talho para saber se as promoções influenciaram.
141	Gestão de Parceiros	Nº de cupões não usados em toda a loja	Cliente recebeu oferta e não usou.

142	Gestão de Parceiros	Rebates sms notificação	Nº de clientes que rebatem sms notificação, até acabar o prazo de utilização.
143	Gestão de Parceiros	Rebates acima média	Clientes que rebateram cupões de mailing no ecossistema e por parceiro, acima da média.
144	Campaign Intelligence	Nuvem clicks nos folhetos	Quais as zonas mais clicadas no folheto promocional. Filtrar por grupos.
145	Campaign Intelligence	Clientes preferenciais feriados	Clientes que preferem fazer compras nos feriados.
146	CCR - Categoria	Segmentação 7	Segmentação 7 por top grupos de produtos.
147	CCR - Operacional	VL produtos em folheto	Saber vendas dos produtos em folheto, saber o impacto do produto no reforço do folheto. Vendas ao artigo.
148	CCR - Operacional	Clientes em redução	Número de rebates nos últimos 12 meses dos clientes com redução de gasto.
149	CCR - Operacional	Brinquedos campanha	Clientes que compraram brinquedos durante a campanha. Clientes com afinidade campanha temática.
150	CCR - Operacional	Média Áreas Frescos	Quantas áreas de frescos o cliente visitou por transação, em média.
151	CCR - Operacional	Áreas Frescos	Quais as áreas de frescos o cliente visitou por transação, em média Top frescos.

152	Advanced Analytics & Insight	Cliente queda de compra	Não fez compras últimos 12 meses. Apoio nas decisões de notificações.
153	Advanced Analytics & Insight	Ultima categoria nova	Última categoria que não é habitual o cliente comprar.
154	Gestão Promocional	VL por oferta	Métrica base de todas as ofertas; por meio de comunicação (TV, radio, sms, folhetos).
155	Gestão Promocional	Cientes artigo TVRadio	Número de clientes que nunca compraram um artigo mas compraram porque está em TV, radio;
156	Gestão Promocional	Transações por oferta	Número de transações por oferta e por meio de comunicação.
157	Gestão Promocional	Descontos por oferta	Número de descontos por meio de comunicação.
158	Gestão Promocional	Meios Comunicação Zona	Impacto dos meios por zona/distrito, por transações e vendas.
159	Campaign Intelligence	Cliente brand scoring	Clientes com engagement por tipo de marca: marca própria. Cliente que compra em todos os parceiros.
160	Campaign Intelligence	Cliente promoção scoring	Clientes com engagement por promoções.

161	CCR - Cliente/ Just do it	Nº de categorias por UN	Quantas categorias por unidade.
162	CCR - Cliente/ Just do it	Categorias por UN	Quais as categorias por UN.
163	CCR - Cliente/ Just do it	Clientes cat/UN acima/abaixo media cesta	Clientes que comprem na categoria/UN acima/abaixo da média de cesta.
164	CCR - Cliente/ Just do it	Clientes cat/UN acima/abaixo media frequência	Clientes que comprem na categoria/UN acima/abaixo da média de frequência.
165	CCR - Cliente/ Just do it	Clientes em redução por loja	Número de clientes em redução por gasto, total ecossistema.
166	CCR - Cliente/ Just do it	Clientes nunca Rebateram cupões	Clientes que nunca rebateram cupões.
167	CCR - Cliente/ Just do it	Clientes com perda de gasto nos parceiros U12M	Clientes em perda.
168	CCR - Cliente/ Just do it	Clientes com perda de gasto no parceiro A U12M	Clientes em perda, lojas parceiro A.
169	CCR - Cliente/ Just do it	Clientes com perda de gasto nos parceiros U24M - U12M	Clientes em perda, loja a loja.
170	CCR - Cliente/ Just do it	Clientes com perda de gasto no parceiro A U24M - U12M	Clientes em perda, lojas parceiro A.
171	CCR - Cliente/ Just do it	Rebates cupões enviados	Número de cupões enviados por estratégia M.

172	CCR - Cliente/ Just do it	Rebates cupões rebatidos	Número de cupões rebatidos por estratégia de mailing.
173	CCR - Cliente/ Just do it	Mailings recebidos	Quantos mailings que o cliente recebeu nos U12M.
174	CCR - Cliente/Just do it	Mailings rebatidos	Quantos mailings que o cliente rebateu pelo menos 1 cupão nos U12M; Total.
175	CCR - Cliente/ Just do it	Mailings rebatidos cupão valor	Quantos mailings que o cliente rebateu pelo menos 1 cupão valor nos U12M.
176	CCR - Cliente/ Just do it	Mailings rebatidos cupão O	Quantos mailings que o cliente rebateu pelo menos 1 cupão O nos U12M.
177	CCR - Cliente/ Just do it	Penetração marcas por tipo de marca	Em todos os tipos de marca no gasto, apenas Direcção Comercial Alimentar.
178	CCR - Cliente/ Just do it	Taxa desconto	Taxa de desconto utilizada.
179	CCR - Cliente/ Just do it	Peso frescos	Percentagem de vendas de frescos.
180	CCR - Cliente/ Just do it	Peso top 10 categoria nas VL	Peso do top 10 da categoria nas VL da categoria.
181	CCR - Cliente	Segmentação 12	Segmentação 12 com os critérios dependentes do parceiro.
182	CCR - Cliente	VL reportadas	Vendas líquidas reportadas;

183	CCR - Cliente	Missão de compra	Número de transações por tipo de missão.
184	Campaign Intelligence	Tipo pagamento preferencial	Pagamento preferencial.
185	Advanced Analytics & Insight	Segmentação 78	Se o cliente acumula o seu saldo no cartão do programa de fidelização ou se o gasta assim que tem oportunidade, para os parceiros A, C e D.
186	Advanced Analytics & Insight	Segmentação 9	Identificar o período preferencial do mês Parceiros A, C e D
187	Campaign Intelligence	Cliente vegetariano	Saber se o cliente compra artigos que o considerem vegetariano.
188	CCR - Cliente	Preferência cupão ou POS	Preferência por cupão ou por POS. Parceiros A, C e D
189	CCR - Cliente	Clientes exclusivo campanha	Clientes exclusivos por campanha temática Parceiros A, C e D
190	CCR - Cliente	Reclamação compra	Flag de reclamação e nova compra.
191	CCR - Cliente	Nível promocional do cliente	Cliente com grau preferencial de desconto -se possível por categoria.
192	CCR - Cliente	Cliente take-out	Sensibilidade do cliente ao take out (produtos perto da caixa de pagamento: chicletes, pilhas...).

193	CCR - Cliente	VL dieta	Vendas líquidas dos produtos da secção dieta.
194	CCR - Cliente	Cliente marca	Saber se o cliente compra marca própria ou fornecedor.
195	CCR - Cliente	Devoluções	Número de devoluções por cliente.
196	Campaign Intelligence	Desconto acumulado rebatido	Por parceiro e total do ecossistema.
197	Campaign Intelligence	Desconto acumulado expirado	Por parceiro e total do ecossistema.
198	Advanced Analytics and Insight	Segmentação 2	Parceiros A, C e D e tem objetivos semelhantes à Segmentação 1 mas para os últimos 12 meses.

ANEXO C: Documento de Levantamento de requisitos

A	B	C	D	E	F	G	H	J	K	L	M
Nº	Equipe	Indicador	Descrição	Fórmula/retorno	Período de cálculo/análise	Período de actualização	Prioridade	Usabilidade (x por semana/mês)	Utilidade	Notas	Existência
1	Campaign Intelligence	Segmentação 1	Segmentação 1 é calculada para os parceiros A, C e D tendo como objetivos compreender o impacto de novos formatos de loja em cada nível de valor e identificar potenciais clientes de alto valor que compram noutros parceiros.	Qualitativa Ordinal string; nome do segmento	segmentação mais atual	Mensal	Priority	1x mês	Ações promocionais		sim
2	Campaign Intelligence	Segmentação 1 parceiros A, C e D U12m	Esta segmentação é calculada para os parceiros A, C e D e tem objetivos semelhantes à Segmentação 1 mas para os últimos 12 meses.	Qualitativa Ordinal string; nome do segmento	segmentação mais atual	Mensal	Priority	1x mês	Ações promocionais		sim
3	Campaign Intelligence	Motivável	Permissão de contactar o cliente via telefone.	Qualitativa Nominal flag	estático (não tem período)	Semanal	Priority	1x mês	Ações promocionais; Reports		sim
4	Campaign Intelligence	Emailável	Permissão de contactar o cliente via email.	Qualitativa Nominal flag	estático (não tem período)	Semanal	Priority	1x mês	Ações promocionais; Reports		sim
5	Campaign Intelligence	Maiável	Permissão de contactar cliente via carta (morada).	Qualitativa Nominal flag	estático (não tem período)	Mensal	Priority	1x mês	Ações promocionais; Reports		sim
6	Campaign Intelligence	POSável	Utilizador do sistema POS.	Qualitativa Nominal flag	estático (não tem período)	Mensal	Priority	1x mês	Ações promocionais		sim
7	Campaign Intelligence	Appável	Utilizador da app parceiro K.	Qualitativa Nominal flag	estático (não tem período) CCR: últimos 12 meses; U24-30; YTD do ano corrente e anterior	Mensal	Priority	1x mês	Ações promocionais		não
8	Campaign Intelligence	Tiravestível	Utilizador do sistema tira-vez.	Qualitativa Nominal flag	estático (não tem período) CCR: últimos 12 meses; U24-30; YTD do ano corrente e anterior	Mensal	Priority	1x mês	Ações promocionais		não
9	Campaign Intelligence	Cliente ativo no parceiro	Permite saber se o cliente fez compras nos últimos 12 meses para cada parceiro.	Qualitativa Nominal flag	últimos 12 meses	Mensal; período considerado cliente período (8 semanas sem registo)	Medium priority	1x mês	Ações promocionais; Reports	Flag + total de compras	não
10	Campaign Intelligence +LTP	Cliente Zona	Cliente Zona apresenta se o cliente é residente em Portugal continental ou nas ilhas (Açores ou Madeira) através da análise da morada registada.	Qualitativa Ordinal string; nome do segmento	se for loja preferencial são 8 semanas	Semanal	Medium priority	1x mês	Ações promocionais	Testrangeiros, migrantes	sim
16	Business Analytics	Segmentação 4	A segmentação 4 permite uma análise por tipologia de produtos nos últimos 12 meses para os parceiros A, C e D.	Qualitativa Ordinal string; nome do segmento	segmentação mais atual	Mensal	Nice to have	1x mês	Análises ad-hoc		nao existe por parceiros
17	Business Analytics	Segmentação 5	A Segmentação 5 tem base no comportamento de compra do cliente, para os parceiros A, C e D retorna um resultado do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo dividido em 3 patamares de afinidade a produtos de crianças.	Qualitativa Ordinal string; nome do segmento	segmentação mais atual	Mensal	Nice to have	1x mês	Análises ad-hoc	TCCR saber se é transferir a segmentação de outra ferramenta	sim
18	Advanced Analytics & Insight	Valor cliente	Valor do cliente a longo prazo; prever quanto o gasto do cliente nos diferentes parceiros.	Qualitativa Ordinal real	últimos 12 meses	Mensal	Nice to have	1x em 3 meses	Desenvolvimento de projetos		não
19	Advanced Analytics & Insight	Nº de sugestões	Nº de sugestões efetuadas.	Quantitativo Numérico int	últimos 12 meses	Mensal	Nice to have	1x ano	Análises ad-hoc		sim
20	Advanced Analytics & Insight	Nº de reclamações	Nº de reclamações efetuadas.	Quantitativo Numérico int	últimos 12 meses	Mensal	Nice to have	1x ano	Análises ad-hoc		sim
21	Gestão de Parceiros	Filhos parceiro K	Em concordância com os dados dos clientes do parceiro K o resultado classifica-se por ser do tipo quantitativo numérico sendo o número de filhos do cliente.	Qualitativa Ordinal string; nome do segmento	segmentação mais atual	Mensal	Priority	1x ano	Ações promocionais		não

ANEXO D: Indicadores da estrutura de dados LAURIUS

Nº	Indicador	Definição
1	Segmentação 1	<p>Segmentação 1 é calculada para os parceiros A, C, D, F, H, J e K tendo como objetivos compreender o impacto de novos formatos de loja em cada nível de valor e identificar potenciais clientes de alto valor que comprem noutros parceiros.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e é atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de ações promocionais.</p>
2	Segmentação 1 parceiros A, C e D U12m	<p>Esta segmentação é calculada para os parceiros A, C e D e tem objetivos semelhantes à Segmentação 1 mas para os últimos 12 meses.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e é atualizada uma vez por mês.</p> <p>A sua utilidade é para ajudar no desenvolvimento de ações promocionais.</p>
3	Segmentação 3	<p>Segmentação 3 é classificada pela frequência do cliente de acordo com o parceiro.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de ações promocionais.</p>
4	Segmentação 4	<p>A segmentação 4 permite uma análise por tipologia de produtos nos últimos 12 meses para os parceiros A, C e D.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de análises ad-hoc.</p>
5	Segmentação 5	<p>A Segmentação 5, com base no comportamento de compra do cliente, para os parceiros A, C e D, retorna um resultado do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo dividido em 3 patamares de afinidade a produtos de crianças.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de análises ad-hoc.</p>

6	Segmentação 6 parceiro A	<p>Segmentação das lojas do parceiro A.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada trimestralmente. A sua utilidade é para ajudar no desenvolvimento de análises ad-hoc.</p>
7	Segmentação 4 top produtos	<p>Esta segmentação respeita parâmetros de admissão internos nos últimos 12 meses para cada parceiro.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de análises ad-hoc.</p>
8	Segmentação 8	<p>Dependendo do parceiro, esta segmentação considera diferentes tipos de categorias.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para apoiar decisões nas ações promocionais.</p>
9	Segmentação 9	<p>A Segmentação 9 permite identificar o período preferencial do mês.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o nome do segmento atribuído ao retorno do cálculo.</p> <p>A segmentação considerada é a mais atual e atualizada uma vez por mês. A sua utilidade é para ajudar no desenvolvimento de ações promocionais.</p>
10	Mobilable	<p>O resultado deste indicador apresenta-se por ser do tipo qualitativo nominal sendo uma flag com dois possíveis retornos definindo se o cliente permite ou não ser contactado via telemóvel.</p> <p>O período de atualização é semanal e tem utilidade para ações promocionais e relatórios.</p>
11	Emailable	<p>O resultado deste indicador é do tipo qualitativo nominal sendo uma flag com dois possíveis retornos definindo se o cliente permite ou não ser contactado via email.</p> <p>O período de atualização é semanal e tem utilidade para ações promocionais e relatórios.</p>
12	Nº de SMS enviados	<p>Apresenta um contador de mensagens enviadas até ao momento por parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico e como atualização mensal. Demonstra utilidade para as análises ad-hoc.</p>
13	Loja preferencial	<p>Loja preferencial por parceiro considerando a segmentação mais atual.</p>

		<p>O resultado deste indicador é do tipo quantitativo numérico, sendo o código da localização da loja preferencial.</p> <p>O período de atualização é mensal sendo útil para análises ad-hoc e contadores.</p>
14	Hora preferencial de compra	<p>Parte do dia preferencial do cliente para fazer compras, calculado por parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo o tempo (manhã, almoço, tarde, fim de tarde) preferencial.</p> <p>O período de atualização é mensal sendo útil para análises ad-hoc.</p>
15	Dia de semana preferencial	<p>Dia da semana preferencial do cliente para fazer compras, calculado por parceiro.</p> <p>O resultado deste indicador é do tipo qualitativo nominal, sendo o dia da semana preferencial.</p> <p>O período de atualização é mensal sendo útil para análises ad-hoc.</p>
16	Parceiro A preferencial	<p>Cliente com loja preferencial dos parceiros A, C e D adjacente às lojas por parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo o código da loja.</p> <p>O período de atualização é mensal sendo útil para ações promocionais.</p>
17	Cartão novo programa fidelização preferencial	<p>Número do novo cartão de fidelização associado ao programa de fidelização da empresa.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo o código do cartão.</p> <p>O período de atualização é diário sendo útil para contadores, análises ad-hoc e ações promocionais.</p>
18	Categoria preferencial	<p>Categoria preferencial do cliente, por parceiro.</p> <p>O resultado deste indicador é do tipo qualitativo nominal, sendo o nome da categoria.</p> <p>O período de atualização é mensal sendo útil para contadores e análises ad-hoc.</p>
19	Distância à loja preferencial	<p>Distância à loja preferencial definida, de acordo com o parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo os metros entre a morada e a loja preferencial.</p> <p>O período de atualização é mensal sendo útil para análises ad-hoc.</p>
20	Distância à loja mais próxima	<p>Distância à loja mais próxima, de acordo com o parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo os metros entre a morada e a loja mais próxima desta.</p> <p>O período de atualização é mensal sendo útil para análises ad-hoc e</p>

		relatórios.
21	Código Postal	<p>Código postal é um conjunto de quatro algarismos seguido de mais três algarismos que precede o nome de uma localidade utilizado para facilitar o encaminhamento e a entrega de correspondência. Neste sentido a base de códigos postais é fornecida pela empresa CTT - Correios de Portugal (Correios, Telégrafos e Telefones) sendo então validado o código postal registado pelo cliente no momento da adesão do cartão de fidelização apenas para códigos em Portugal.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, sendo o código postal da morada do cliente.</p> <p>O período de atualização é mensal sendo útil para ações promocionais.</p>
22	Cliente ativo no parceiro	<p>Permite saber se o cliente fez compras nos últimos 12 meses para cada parceiro.</p> <p>O resultado deste indicador é do tipo qualitativo nominal, sendo uma flag correspondendo a cliente ativo ou não.</p> <p>O período de atualização é consoante o período considerado um cliente perdido, sendo útil para ações promocionais e relatórios.</p>
23	Cliente Zona	<p>Cliente Zona apresenta se o cliente é residente em Portugal continental ou nas ilhas (Açores ou Madeira) através da análise da morada registada.</p> <p>O resultado deste indicador é do tipo qualitativo nominal, definido como possibilidades Portugal, Madeira ou Açores.</p> <p>O período de atualização é semanal, sendo útil para ações promocionais.</p>
24	Promoção W	<p>Cliente que adere à promoção W.</p> <p>O resultado deste indicador é do tipo qualitativo nominal, definido como uma flag se adere à promoção em questão.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais e crédito na promoção primavera.</p>
25	Exclusivo parceiro	<p>Valor de vendas por parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico, definido como o valor em vendas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
26	Cliente do novo cartão de fidelização	<p>Cliente portador do novo cartão de fidelização.</p> <p>O resultado deste indicador classifica-se por ser do tipo qualitativo nominal sendo uma flag acionada se a conta cliente for uma conta do novo cartão de fidelização.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e</p>

		ações promocionais.
27	Plano de parceria Utilities	<p>Cliente aderente ao plano de parceria Utilities.</p> <p>O resultado deste indicador classifica-se por qualitativo nominal sendo uma flag acionada caso o cliente adira ao plano.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc, ações promocionais e relatórios.</p>
28	Balanço	<p>O resultado deste indicador classifica-se por tipo qualitativo nominal sendo uma flag acionada para o cliente propenso ao abandono.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
29	Cliente brand	<p>Cliente com compromisso por marca, realizando compras em todos os parceiros.</p> <p>O resultado deste indicador classifica-se por qualitativo nominal sendo uma flag acionada para o cliente com compromisso por marca.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
30	Cliente P	<p>O resultado deste indicador classifica-se por qualitativo nominal sendo uma flag acionada quando o cliente é uma conta P.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
31	Cliente L	<p>O resultado deste indicador classifica-se por qualitativo nominal sendo uma flag acionada quando o cliente tem uma conta L.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
32	Estado da conta	<p>Estado da conta do cliente conforme os critérios internos definidos.</p> <p>O resultado deste indicador classifica-se por qualitativo nominal sendo o nome do estado da conta.</p> <p>O período de atualização é semanal, sendo útil para ações promocionais.</p>
33	Nº de rebatesU12M	<p>Contador de rebates, nomeadamente utilização de cupão de benefícios para o cliente, calculado para cada parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo número de rebates.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
34	Nº de não rebates	<p>Contador de rebates em que os cupões registados em caixa não foram rebatidos, para cada parceiro.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo</p>

		<p>número de não rebates.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
35	Nº de rebates U12M, por estratégia	<p>Contador de rebates nos últimos 12 meses em conformidade com as estratégias definidas para o parceiro A. Contador também calculado para os restantes parceiros separando por outra estratégia diferente das dos parceiros A, C e D.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo número rebates.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
36	Rebates sms notificação	<p>Percentagem associada aos rebates referentes às mensagens enviadas de notificação de utilização de um determinado desconto ou promoção.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor da percentagem.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e ações promocionais.</p>
37	Desconto acumulado U12M	<p>Desconto acumulado do cliente nos últimos 12 meses por parceiro e total do ecossistema.</p> <p>O resultado deste indicador classifica-se por tipo quantitativo numérico sendo o valor acumulado dos descontos.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
38	Desconto acumulado Umes	<p>Desconto acumulado do cliente no último mês definido como concluído nos termos internos de finalização do mês. Calculado por parceiro incluindo o serviço on-line e total do ecossistema.</p> <p>O resultado deste indicador classifica-se por quantitativo numérico sendo o valor acumulado dos descontos.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
38	Desconto de preço obtido U12M	<p>Desconto direto que o cliente obteve no preço nos últimos 12 meses, calculado para cada parceiro e total do ecossistema.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor acumulado do desconto obtido.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
40	Desconto de preço obtido Umes	<p>Desconto direto que o cliente obteve no preço no último mês, calculado para cada parceiro e total do ecossistema.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor acumulado do desconto obtido.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e</p>

		contadores.
41	Taxa desconto	<p>Taxa de desconto direto e em cartão.</p> <p>O resultado deste indicador é tipo quantitativo numérico sendo o valor da taxa.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
42	Vendas brutas U12M	<p>Vendas brutas do cliente nos últimos 12 meses por parceiro e total do ecossistema. Por outras palavras, o valor original de vendas sem descontos associados.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas brutas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
43	Vendas brutas Umes	<p>Vendas brutas do cliente no último mês por parceiro e total do ecossistema. Por outras palavras, o valor original de vendas sem descontos associados.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas brutas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
44	Vendas Líquidas U12M	<p>Vendas líquidas do cliente nos últimos 12 meses por parceiro e total do ecossistema. Por outras palavras, o valor registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para ad-hoc e contadores.</p>
45	Vendas Líquidas Umes	<p>Vendas líquidas do cliente no último mês por parceiro e total do ecossistema. Por outras palavras, o valor registado no último mês.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
46	VL DC Alimentar	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial alimentar, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A, registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>

47	VL DC Frescos	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial frescos, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
48	VL DC Têxtil	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial têxtil, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
49	VL DC Bazar	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial bazar, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
50	VL DC Casa	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial casa, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
51	VL DC Foods & Bakery	<p>Vendas líquidas do cliente nos últimos 12 meses na direção comercial Foods and Bakery, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e contadores.</p>
52	VL online U12M	<p>Vendas líquidas do cliente nos últimos 12 meses no serviço on-line, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p>

		O período de atualização é mensal, sendo útil para análises ad-hoc.
53	VL online Umes	<p>Vendas líquidas do cliente no último mês no serviço on-line, para o parceiro A. Designadamente o valor referente ao parceiro A registado no último mês.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
54	VB online U12M	<p>Vendas brutas do cliente nos últimos 12 meses no serviço on-line, para o parceiro A. Designadamente o valor referente ao parceiro A, registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas brutas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
55	VB online Umes	<p>Vendas brutas do cliente no último mês no serviço on-line, para o parceiro A. Designadamente o valor referente ao parceiro A, registado no último mês.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas brutas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
56	VL dieta	<p>Vendas líquidas do cliente na categoria de diatéticos nos últimos 12 meses, para os parceiros A, C e D. Designadamente o valor referente ao parceiro A registado nos últimos 12 meses.</p> <p>O resultado deste indicador é do tipo quantitativo numérico sendo o valor das vendas líquidas.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
57	Filhos parceiro K	<p>Em concordância com os dados dos clientes do parceiro K o resultado é do tipo quantitativo numérico sendo o número de filhos do cliente.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e relatórios.</p>
58	Plano Y parceiro H	<p>Cliente elegível ou se já utilizou o plano Y do parceiro H.</p> <p>O resultado é do tipo qualitativo nominal sendo a flag acionada quando o cliente é elegível do plano Y do parceiro H.</p> <p>O período de atualização é diário, sendo útil para ações promocionais como o envio de mensagens ou newsletters.</p>
59	Nº de lojas frequentadas	<p>Contador de lojas, por parceiro, frequentadas pelo cliente.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número de lojas.</p>

		O período de atualização é mensal, sendo útil para ações promocionais, contadores e análises ad-hoc.
60	Transações U12M	<p>Contador de visitas do cliente nos últimos 12 meses, por parceiro e total do ecossistema.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número de visitas do cliente.</p> <p>O período de atualização é mensal, sendo útil para contadores e análises ad-hoc.</p>
61	Transações Umes	<p>Contador de visitas do cliente no último mês, por parceiro e total do ecossistema.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número de visitas do cliente.</p> <p>O período de atualização é mensal, sendo útil para contadores e análises ad-hoc.</p>
62	Saldo disponível	<p>Saldo disponível no cartão do program de fidelização.</p> <p>O resultado classifica-se por quantitativo numérico sendo o valor do saldo no cartão do programa de fidelização.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
63	Género	<p>Género que identifica e diferencia os homens e as mulheres.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o retorno do cálculo possibilitar 3 tipos feminino, masculino ou nulo quando não tem valor atribuído ou possivelmente o cliente optou por não indicar.</p> <p>A sua utilidade é para ajudar no desenvolvimento de ações promocionais e é atualizado mensalmente.</p>
64	Data de nascimento	<p>Data de nascimento identifica o dia em que o cliente nasceu sendo posteriormente possível calcular a sua faixa etária.</p> <p>O resultado deste indicador é do tipo qualitativo ordinal, sendo o retorno do cálculo possibilitar 2 tipos ano.mês.dia ou nulo quando não tem valor atribuído.</p> <p>A sua utilidade é no apoio do desenvolvimento de ações promocionais sendo atualizado mensalmente.</p>
65	Correspondência devolvida	<p>Este indicador viabiliza a identificação de moradas inválidas ou com erros.</p> <p>O resultado classifica-se por qualitativo nominal sendo a flag acionada quando a morada registada é inválida.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
66	Recency	<p>Número exato de dias que o cliente não fez compras, por parceiro.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número</p>

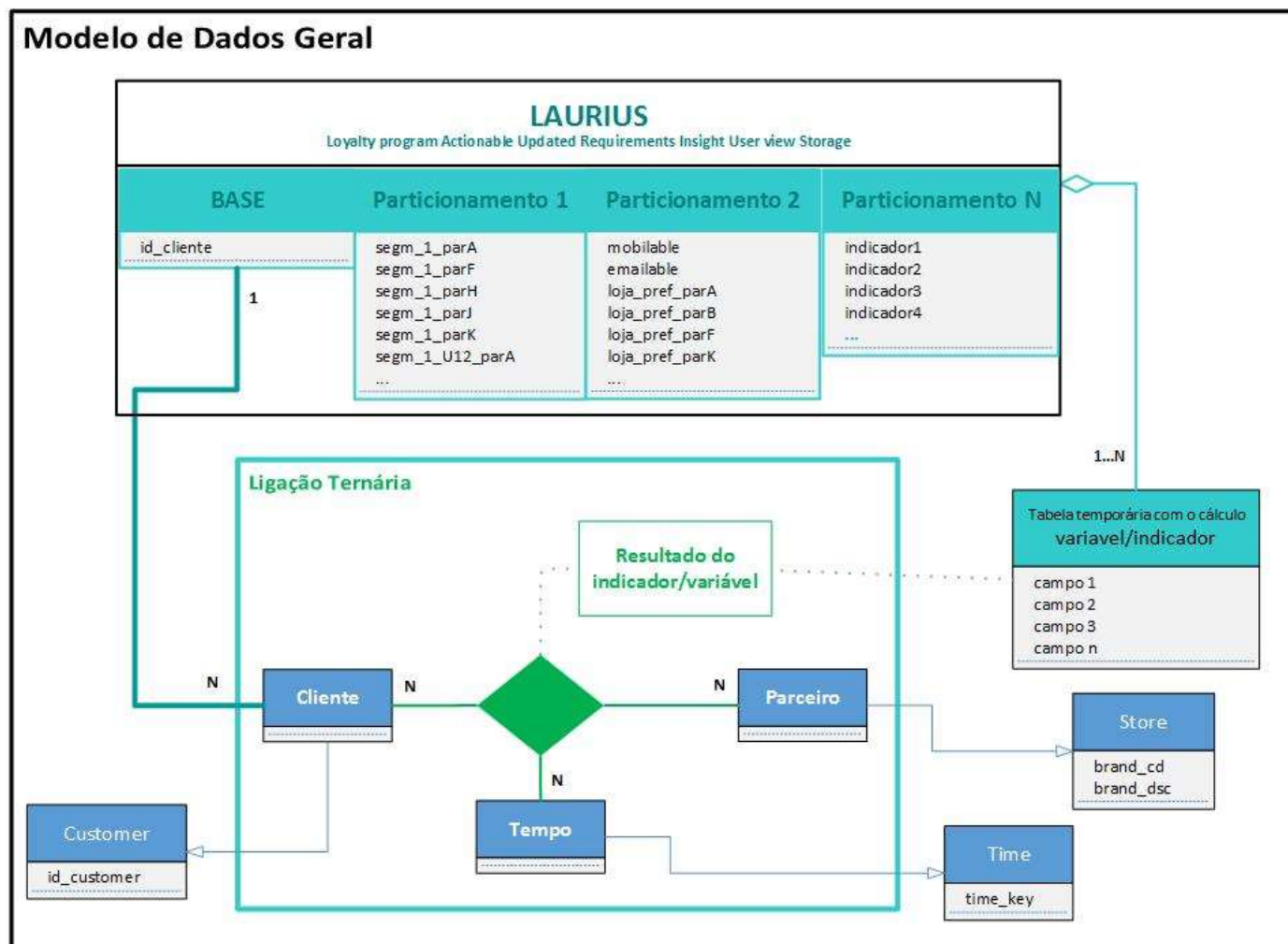
		<p>de dias que o cliente não fez compras.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc e relatórios.</p>
67	Nº de produtos distintos no parceiro A	<p>Contador de produtos diferentes que o cliente comprou nos últimos 12 meses nos parceiros A, C e D.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número produtos distintos comprados.</p> <p>O período de atualização é mensal, sendo útil para o modelo de cálculo de scores.</p>
68	Nº de reclamações	<p>Contador de reclamações efetuadas pelo cliente nos últimos 12 meses.</p> <p>O resultado classifica-se por quantitativo numérico sendo o número de reclamações registadas.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc.</p>
69	Self-check-out	<p>Contador de transações de self-check-out nos últimos 12 meses no parceiro A.</p> <p>O resultado classifica-se por quantitativo numérico sendo o resultado do cálculo definido.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc, contadores e ações promocionais.</p>
70	Utilizações Online	<p>Contador de utilização do serviço on-line nos últimos 12 meses, nos parceiros A,C, D, H, J e K.</p> <p>O resultado classifica-se por qualitativo nominal sendo a flag acionada quando o cliente já utilizou o serviço on-line no parceiro indicado.</p> <p>O período de atualização é mensal, sendo útil para análises ad-hoc, contadores e ações promocionais.</p>
71	Missão de compra	<p>Contador de transações nos últimos 12 meses por tipo de missão no parceiro A, definida pelos termos internos da empresa.</p> <p>O resultado é do tipo quantitativo numérico sendo o número de transações por tipo de missão de compra.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>
72	Devoluções	<p>Número de devoluções do cliente nos últimos 12 meses, nos parceiros A, C e D.</p> <p>O resultado é do tipo quantitativo numérico sendo o número de devoluções registadas.</p> <p>O período de atualização é mensal, sendo útil para ações promocionais.</p>

73	Mailable	<p>Este indicador encontra-se em fase de avaliação de critérios internos do seu cálculo e intuito.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal sendo uma flag acionada caso o cliente permita o contacto via carta.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais e relatórios.</p>
74	Gasto parceiro A	<p>Este indicador encontra-se em fase de avaliação de critérios internos do seu cálculo e intuito.</p> <p>O resultado deste indicador poderá ser do tipo quantitativo numérico sendo a percentagem gasta em orçamento nos últimos 12 meses no parceiro A.</p> <p>O período de atualização poderá ser mensal com utilidade para análises ad-hoc.</p>
75	Valor cliente	<p>Este indicador encontra-se em fase de avaliação de critérios internos do seu cálculo e intuito.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo ordinal sendo o valor do cliente a longo prazo para prever quanto o cliente irá gastar nos diferentes parceiros.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais.</p>
76	Cliente promoção scoring	<p>Este indicador encontra-se em fase de avaliação de critérios internos do seu cálculo e intuito.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal sendo a flag acionada quando o cliente tem um compromisso com promoções ou do tipo quantitativo numérico sendo a percentagem do compromisso do cliente pelas promoções.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais.</p>
77	Contacto preferencial - channel scoring	<p>Este indicador encontra-se em fase de avaliação de critérios internos do seu cálculo e intuito.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal sendo o meio de contato preferencial do cliente.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais e análises ad-hoc.</p>
78	Segmentação 78	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois é uma segmentação que calcula se o cliente acumula o seu saldo no cartão do programa de fidelização ou se o gasta assim que tem oportunidade, para os parceiros A, C e D.</p>

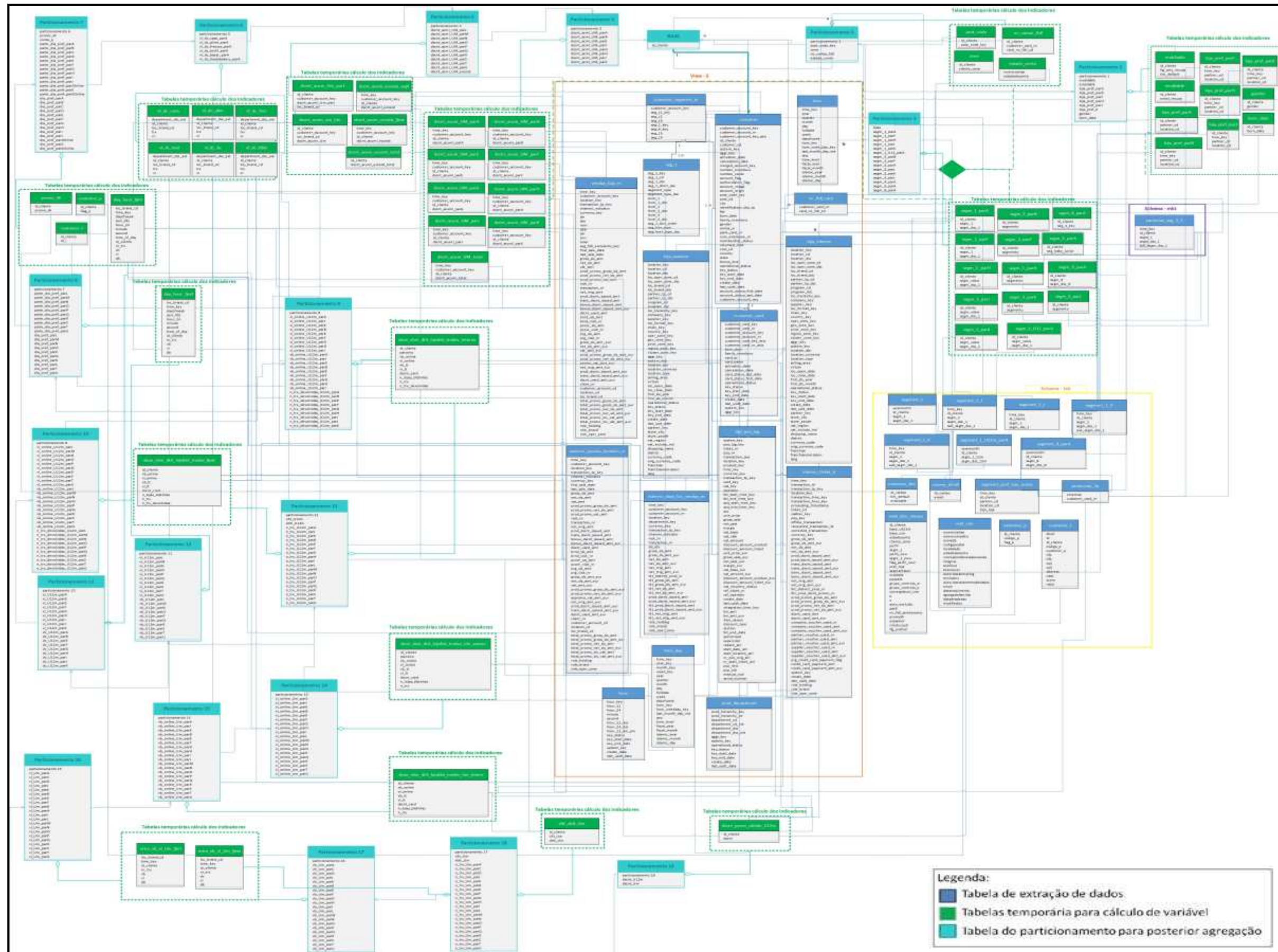
		<p>O resultado deste indicador poderá ser do tipo qualitativo nominal sendo o meio de contato preferencial do cliente.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais e análises ad-hoc.</p>
79	Appable	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois é um meio relativamente recente o que poderá permitir identificar clientes que utilizaram a aplicação móvel do programa de fidelização.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal, sendo a flag acionada caso o cliente tenha utilizado a aplicação nos últimos 12 meses.</p> <p>O período de atualização poderá ser mensal com utilidade para ações promocionais.</p>
80	Weather preferencial	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois indicará o tempo preferencial do cliente, a nível meteorológico na loja de cada parceiro. Pode ser classificado como um indicador com um elevado grau de dificuldade em termos de extração e processamento de dados.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal, sendo o nome do elemento meteorológico preferencial do cliente, por exemplo chuva.</p> <p>O período de atualização poderá ser mensal com utilidade para análises ad-hoc e relatórios.</p>
81	Nuvem clicks folhetos	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois indicará as zonas mais clicadas no folheto promocional disponibilizado on-line. Poderá haver um critério de filtro por grupo, por exemplo considerando a faixa etária de 18 a 30 anos de idade, que irá analisar quais as categorias que foram registadas através do número de cliques do cliente registados.</p> <p>O resultado deste indicador poderá ser do tipo qualitativo nominal, sendo o nome das zonas ou das categorias do folheto.</p> <p>O período de atualização poderá ser mensal com utilidade para análises ad-hoc e ações promocionais.</p>
82	Clicks e-mail	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois indicará se o cliente clicou na hiperligação enviada juntamente com o <i>e-mail</i>, podendo assim revelar o interesse deste.</p>

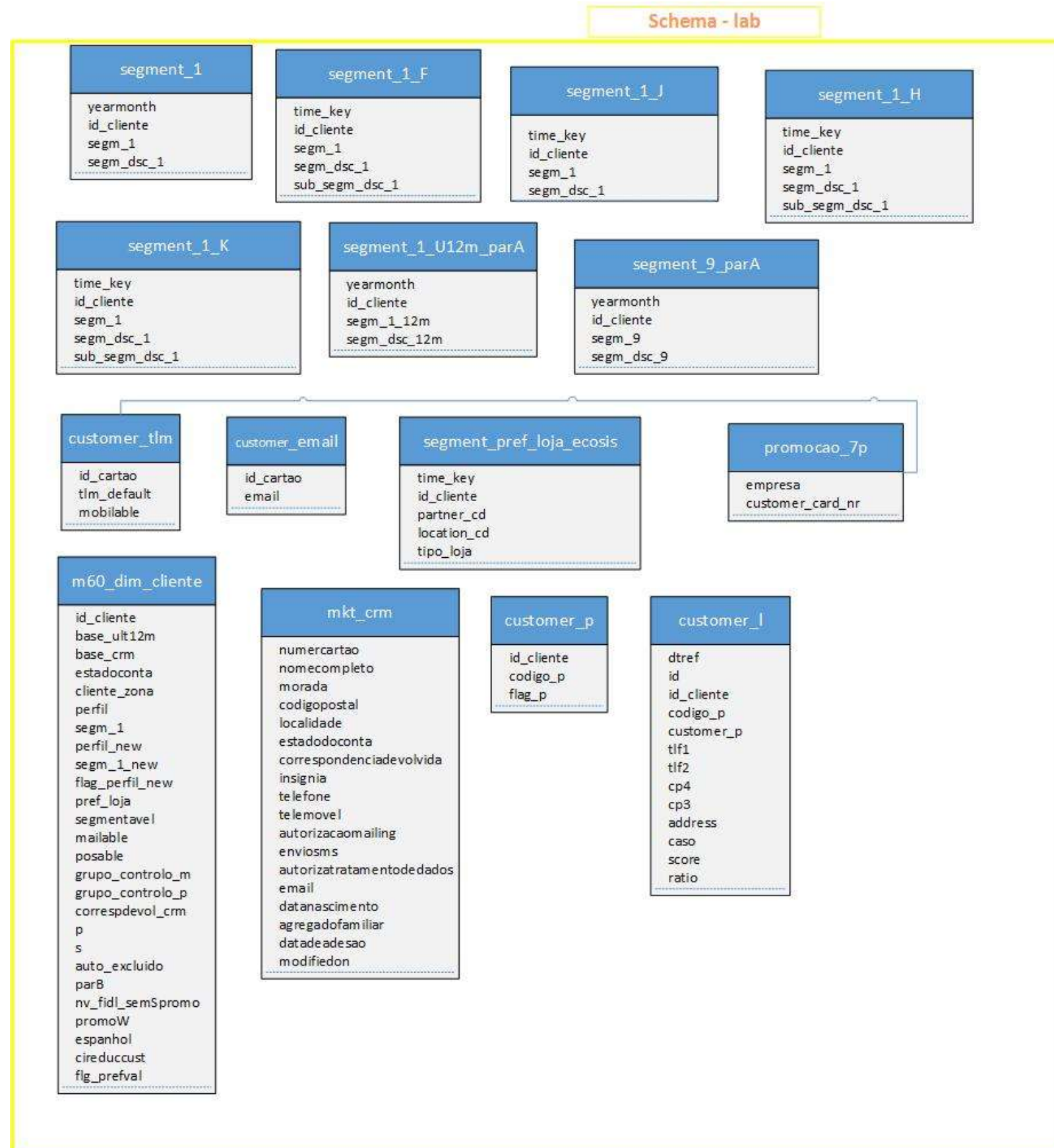
		<p>O resultado deste indicador poderá ser do tipo qualitativo nominal, sendo a flag acionada caso o cliente clique na hiperligação ou do tipo quantitativo numérico sendo o número de vezes que o cliente visitou a hiperligação nos últimos 12 meses comparando com o prazo final da campanha ou promoção enviada no email.</p> <p>O período de atualização poderá ser mensal ou diário dependendo do número de emails enviados com utilidade para análises ad-hoc e ações promocionais.</p>
83	Desconto acumulado expirado	<p>Este indicador não existe e foi proposto na fase de levantamento de requisitos das equipas.</p> <p>É considerado ‘fora da caixa’ pois indicará o valor do desconto total acumulado que o cliente tinha disponível mas que não foi utilizado expirando depois o seu prazo de utilização.</p> <p>O resultado deste indicador poderá ser do tipo quantitativo numérico sendo valor do desconto acumulado expirado.</p> <p>O período de atualização poderá ser mensal ou semanal com utilidade para análises ad-hoc e contadores.</p>

ANEXO E: Modelo de dados geral do projeto LAURIUS

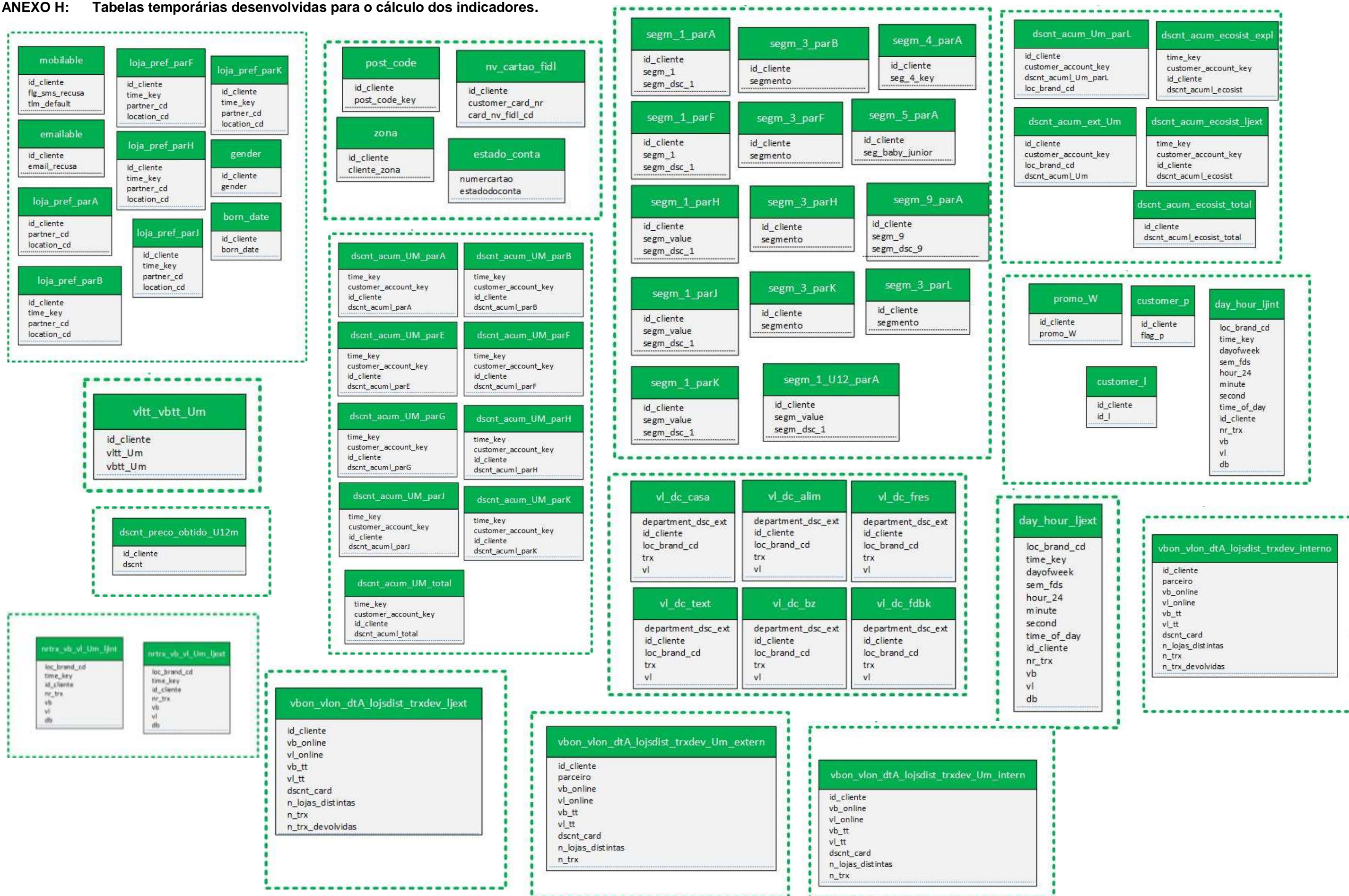


ANEXO F: Modelo de dados detalhado do projeto LAURIUS



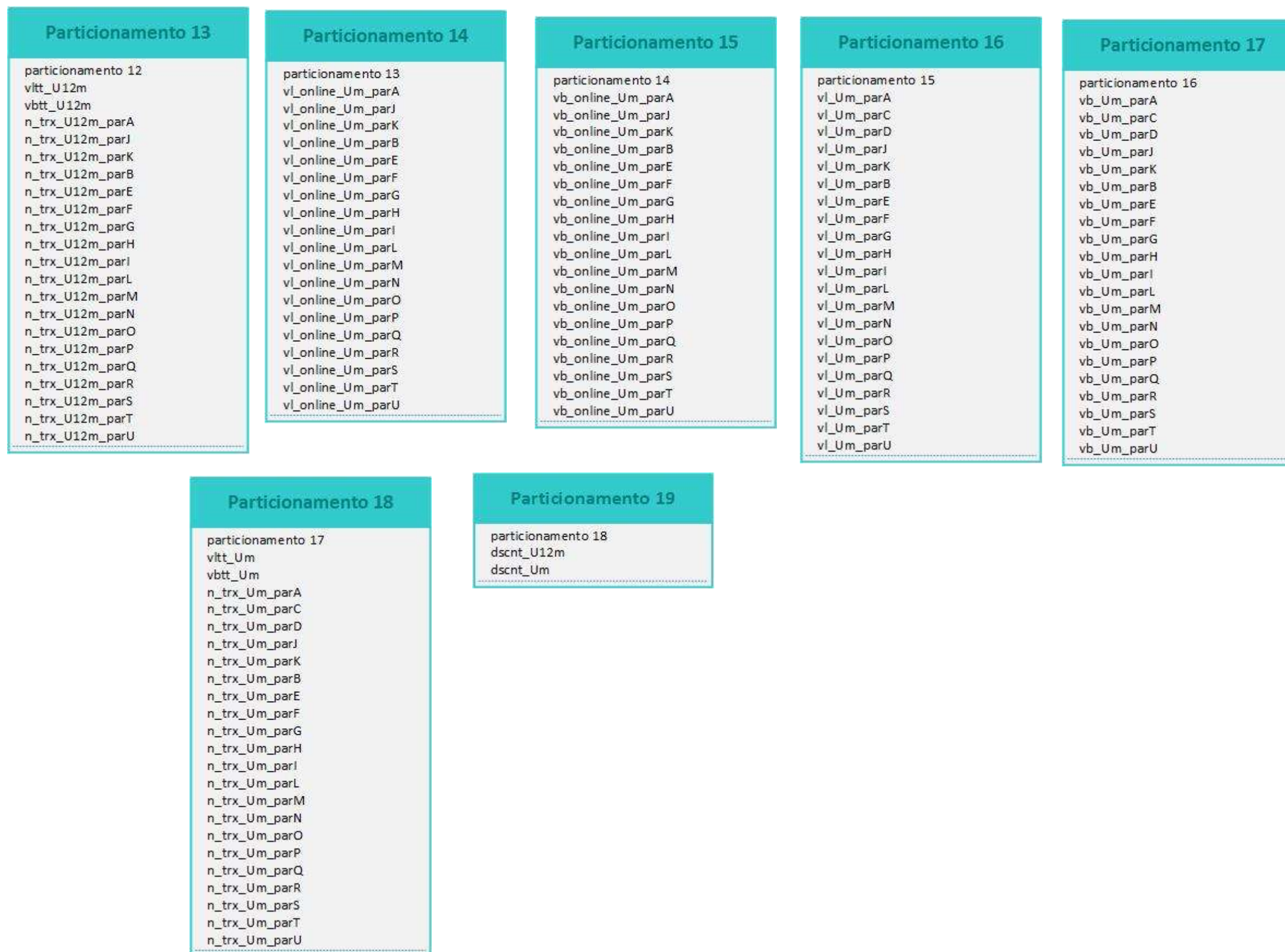
ANEXO G: Tabelas utilizadas dos Schemas lab e mkt e da View E, que pertencem à base de dados analítica.

ANEXO H: Tabelas temporárias desenvolvidas para o cálculo dos indicadores.



ANEXO I: Tabelas das partições implementadas.





ANEXO J: Amostra tabela LAURIUS

LAURIUS_																	
Filter and Sort Query Builder Data Describe Graph Analyze Export Send To																	
	ID_CLIENTE	segm_1_parA	segm_1_parF	segm_1_parH	segm_1_parJ	segm_1_parK	segm_1_U12_paA	segm_3_parB	segm_3_parF	segm_3_parH	segm_3_parJ	segm_3_parK	segm_3_parL	segm_4_parA	segm_5_parA		
1	1	2	5	99	52	4	62	4	1	1	5	5	2	2	4TS01001	BJ_41	
2	1	4	1	99	30	99	99	1	99	99	5	2	1	2	4KB01001	BJ_13	
3	1	7	2	99	30	3	22	1	1	1	5	4	4	2	4KB01001	BJ_13	
4	1	9	99	99	20	99	99	7	99	99	4	2	99	6	4KB01001	BJ_43	
5	1	0	4	99	99	99	99	4	99	99	99	99	99	6	4BF01001	NA	
6	1	3	2	99	99	99	99	2	1	1	4	99	99	2	4wH01001	NA	
7	1	5	1	1	53	2	62	1	1	5	5	5	5	2	4BF01001	BJ_41	
8	1	6	1	1	10	4	76	2	1	5	5	5	2	2	4wH01001	BJ_12	
9	1	9	1	5	20	99	99	1	1	3	5	2	99	1	4AA01001	NA	
10	1	7	2	99	52	99	99	3	1	99	5	99	99	2	4AA01001	NA	
11	1	0	7	99	99	99	99	6	99	99	1	1	99	6	4AA01001	NA	
12	1	3	1	99	99	99	99	3	1	99	99	99	99	5	4KFC01001	-1	
13	1	0	1	99	99	4	76	1	1	1	2	4	2	2	4AA01001	BJ_14	
14	1	1	99	99	99	99	99	5	99	99	1	1	99	6	4AA01001	-1	
15	1	3	1	99	54	2	99	2	99	99	4	4	1	3	4TS01001	BJ_43	
16	1	4	3	99	20	99	99	5	2	1	5	99	99	6	4BF01001	NA	
17	1	5	5	7	99	99	99	3	99	99	1	99	99	6	4KFC01001	NA	
18	1	7	3	99	52	8	99	1	1	99	4	2	1	6	4wH01001	BJ_43	
19	1	8	4	99	99	99	99	8	99	99	1	99	99	6	4KFC01001	-1	
20	1	9	4	7	99	99	99	4	4	99	2	3	99	6	4wH01001	NA	
21	1	9	5	99	99	99	99	3	1	99	1	99	1	4	4AA01001	BJ_43	
22	1	1	1	99	20	99	12	1	1	4	5	99	5	2	4wH01001	BJ_41	
23	1	2	5	99	99	99	99	2	1	99	2	99	99	5	4AA01001	NA	
24	1	3	3	99	30	3	62	5	1	1	5	5	2	6	4AA01001	NA	
25	1	4	99	99	99	99	99	3	1	99	1	99	99	6	4wH01001	NA	
26	1	6	7	99	99	99	99	7	1	99	99	99	99	5	4TS01001	-1	
27	1	6	4	99	99	99	99	4	1	1	2	99	99	2	4TS01001	-1	
28	1	7	3	99	54	1	99	3	99	99	5	5	99	2	4AA01001	BJ_42	
29	1	8	3	5	99	99	99	3	1	99	3	1	99	4	4AA01001	BJ_41	
30	1	9	2	99	54	5	99	1	1	99	2	5	1	5	4BF01001	BJ_41	
31	1	0	3	99	99	99	99	8	99	99	99	99	99	99	4NV01001	NA	
32	1	1	6	99	99	99	99	4	99	99	2	1	99	5	4KFC01001	NA	
33	1	1	4	99	99	99	99	4	1	99	3	99	99	2	4wH01001	NA	
34	1	2	7	99	99	99	99	6	99	99	1	3	99	6	4AA01001	-1	
35	1	3	4	99	99	5	99	1	99	99	1	5	99	2	4AA01001	NA	
36	1	1	2	5	99	99	99	4	1	4	1	99	99	6	4AA01001	NA	
37	1	2	5	99	52	99	99	3	99	99	5	1	99	6	4BF01001	NA	

